

**Determining HDAC8 substrate specificity**

**by**

**Noah Ariel Wolfson**

**A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Biological Chemistry)  
in the University of Michigan  
2014**

**Doctoral Committee:**

**Professor Carol A. Fierke, Chair**

**Professor Robert S. Fuller**

**Professor Anna K. Mapp**

**Associate Professor Patrick J. O'Brien**

**Associate Professor Raymond C. Trievel**

## **Dedication**

My thesis is dedicated to all my family, mentors, and friends who made getting to this point possible.

## Table of Contents

Dedication.....	ii
List of Figures.....	viii
List of Tables.....	x
List of Appendices.....	xi
Abstract.....	xii
Chapter 1 HDAC8 substrates: Histones and beyond.....	1
Overview.....	1
HDAC introduction.....	2
Known HDAC8 substrates.....	4
Candidate non-histone HDAC8 Substrates.....	8
HDAC8 complex formation.....	11
Enzyme structure affects substrate specificity.....	15
Catalytic Mechanism and Regulation of HDAC8 activity.....	22
HDAC8 localization.....	26
Post-translational modification of HDAC8.....	28
Concluding remarks.....	32
Acknowledgements.....	33
Bibliography.....	33
Chapter 2 An enzyme-coupled assay measuring acetate production for profiling histone deacetylase specificity.....	45
Introduction.....	45
Materials and Methods.....	47
Reagents.....	47
Acetyl-CoA synthetase preparation.....	47
HDAC8 expression and purification.....	49

Fluor de Lys assay .....	49
Acetate assay kit .....	50
Optimized stopped coupled acetate assay .....	50
Optimized continuous coupled acetate assay .....	52
Fluorescamine assay .....	52
Results .....	53
Assay .....	53
Optimization of Assay for Measuring HDAC Activity .....	56
Stopped Assay .....	58
Continuous Assay .....	62
Reactivity of HDAC8 with peptides.....	66
Discussion .....	71
Acknowledgements .....	76
Bibliography.....	76
Chapter 3 Profiling small peptide substrates to identify HDAC8 recognition motif(s) .....	81
Introduction .....	81
Materials and Methods .....	82
Reagents.....	82
HDAC8 expression and assay .....	82
Improving the assay conditions .....	83
FlexPepBind protocol .....	84
Modeling the structures of peptide-protein complexes .....	84
Step 1: Preparation of a coarse starting model of a peptide-HDAC8 complex.....	85
Step 2: Optimization of the peptide-HDAC8 complex structure .....	86
Step 2(a) Extensive optimization using Rosetta FlexPepDock .....	86
Step 2(b) Short minimization of starting structure .....	87
Parameters optimized for HDAC8 FlexPepBind.....	87

Definition of axes for rigid body movements.....	89
Addition of an electrostatic term to default scoring function score <sup>12</sup> .....	90
Step 3: Ranking of the peptide substrates based on the optimized peptide-receptor complex structures.....	90
Dataset 1: Training and test set: 361 GXXKacZGC peptides.....	91
Initial calibration set: 5 substrates & 5 non-substrates .....	92
Dataset 2: Validation set: 361 GRKacXZC peptides .....	92
Statistical tests .....	92
Binary distinction between substrates and non-substrates .....	93
Prediction of relative substrate activity .....	93
Results .....	93
First round of calibration .....	93
Analysis of performance in Training and Test sets .....	94
Searching for novel, non-histone substrates .....	95
Experimental validation of de-novo predictions .....	96
Second round of calibration.....	100
Improved starting structure and score function .....	101
Second round of Phosphosite screen .....	103
Assessing the second set of peptides .....	104
Discussion .....	107
Bibliography.....	110
Chapter 4 Mining the Acetylome: An Unbiased Approach to Identify the Endogenous Substrates of Histone Deacetylase 8 (HDAC8) .....	112
Introduction .....	112
Results and discussion.....	112
Acknowledgments.....	125
Methods.....	126
Reagents.....	126

HDAC8 expression and purification .....	126
Peptides.....	128
Peptide deacetylation assay .....	128
Fluorescamine assay .....	129
Compound Synthesis.....	129
Caliper Assay.....	130
Molecular Docking.....	130
Cell Culture and Compound Treatment for Proteomics Experiments.....	131
Proteomics .....	132
Cell Culture and Compound Treatment for Gene Expression Experiments .....	134
Gene Expression.....	135
CMap Query Integration Analysis.....	135
Western Blotting.....	136
Bibliography.....	136
Chapter 5 Acetylated protein substrates have enhanced reactivity with HDAC8 compared to peptide substrates.....	141
Introduction.....	141
Materials and Methods .....	143
Reagents.....	143
HDAC8 expression and purification .....	144
HDAC assay .....	145
Histone expression and purification .....	145
Protein deacetylation assays .....	147
Results.....	147
Deacetylation of acetylated H3/H4 tetramers.....	150
Single turnover kinetics.....	151
Deacetylation of histone octamers.....	154
Multiple turnover of tetramer .....	155

Discussion .....	159
Bibliography .....	162
Chapter 6 Conclusions .....	167
Differentiating substrates from binding partners and downstream effectors .....	167
Identifying the local HDAC-substrate contacts using peptides .....	169
Utilizing proteins to identify long range HDAC-substrate interactions.....	171
Does HDAC8 function in protein complexes .....	172
Bibliography .....	173
Appendices.....	176
Kinetics of peptide substrates tested .....	177
Protocols:.....	180
Coupled acetate assay stopped .....	180
Histone deacetylation mass spectrometry assay (short protocol).....	183
Histone deacetylation mass spectrometry assay (long protocol).....	185
Measuring peptide concentration with Fluorescamine .....	187
HDAC8 purification .....	188
ACS purification.....	192
Acetylated tetramer purification and assembly .....	195
Purifying 601 DNA for nucleosome.....	201

## List of Figures

Figure 1.1 Widely used HDAC inhibitors .....	3
Figure 1.2 The Fluor de Lys assay.....	6
Figure 1.3 HDAC8 structures .....	16
Figure 1.4 HDAC8 with two bound TSA molecules.....	17
Figure 1.5 Structural variation within the structurally characterized HDACs.....	21
Figure 1.6 Schematic of the one base mechanism for HDAC8 .....	23
Figure 1.7 Schematic of three potential models for describing the effect of an activating effector on HDAC activity .....	26
Figure 1.8 Phosphorylation of Ser39 may affect the active site structure and/or reactivity of HDAC8 .....	30
Figure 2.1 HDAC Assay scheme.....	54
Figure 2.2 Comparison of R-biopharm kit and optimized assay for detection of acetate .....	55
Figure 2.3 Extended acetate standard curve .....	56
Figure 2.4 Standard curve for the acetate assay compared to NADH fluorescence.....	58
Figure 2.5 Time course for the formation of acetate .....	60
Figure 2.6 HDAC3/NCOR1 stopped assay time course.....	62
Figure 2.7 Mg <sup>2+</sup> dependence of HDAC8 and acetate assay.....	64
Figure 2.8 Continuous assay to measure HDAC8 activity .....	65
Figure 2.9 HDAC8-catalyzed deacetylation of pan-acetylated H3/H4 histone tetramer.....	66
Figure 2.10 Dependence of HDAC8-catalyzed deacetylation on the concentration of H4 K16ac peptide mimics .....	70
Figure 3.1 Key interactions within the HDAC8-substrate interface.....	89
Figure 3.2 Log(p-value) of KS test vs. activity level plot shows that the best cutoff for binary classification is around 0.34.....	95
Figure 3.3 Distribution of peptide scores in the 3184 acetylated lysine sequences.....	96
Figure 3.4 Experimental validation of the Pep20 set.....	97
Figure 3.5 Improved starting structures for optimization .....	102
Figure 3.6 Distribution of interface score in the acetylated peptides .....	103
Figure 3.7 Performance of new protocol on the second validation set, Pep19.....	107
Figure 4.1 Chemical tools for studying HDAC8 .....	115
Figure 4.2 Scheme and results of mass spectrometry identification of HDAC8 substrates .....	118



Figure 4.3 Deacetylation of peptides identified by mass spectrometry .....	121
Figure 4.4 In vitro peptide deacetylation using commercially available HDACs .....	122
Figure 4.5 The tumor suppressor p21 is linked to HDAC8 inhibition.....	124
Figure 5.1 Schematic of acetylation sites tested .....	148
Figure 5.2 Example substrate dependence for the 13-mer H3K9ac peptide.....	150
Figure 5.3 H3 Histones are singly acetylated .....	151
Figure 5.4 Sample single turnover deacetylation data .....	153
Figure 5.5 H3K14ac octamer data .....	155
Figure 5.6 HDAC8 turnovers vs enzyme concentration.....	156
Figure 5.7 No product inhibition measured .....	158
Figure 5.8 Schematic of inhibition mechanism .....	159

## List of Tables

Table 2.1 Structure and kinetic constants for HDAC8-catalyzed deacetylation .....	69
Table 3.1 Different constraints derived from the solved structure .....	88
Table 3.2 Initial calibration set .....	92
Table 3.3 Distinction between substrates and non-substrates in the initial calibration run.....	94
Table 3.4 Experimental measurement of potential substrates and a corresponding predicted negative set.....	100
Table 3.5 The second experimental validation: Pep19 set.....	106
Table 5.1 Catalytic efficiencies of substrates .....	149

## List of Appendices

Appendices.....	176
Appendix A Kinetics of peptide substrates tested.....	177
Appendix B Protocols .....	180
Coupled acetate assay stopped .....	180
Histone deacetylation mass spectrometry assay (short protocol).....	183
Histone deacetylation mass spectrometry assay (long protocol).....	185
Measuring peptide concentration with Fluorescamine.....	187
HDAC8 purification .....	188
ACS purification.....	192
Acetylated tetramer purification and assembly .....	195
Purifying 601 DNA for nucleosome.....	201

## Abstract

Histone deacetylases (HDACs) are a group of 18 enzymes that catalyze the deacetylation of acetyl lysine residues in proteins. Acetyl lysine residues are present within thousands of proteins, and acetylation/deacetylation has been shown to affect protein properties integral to cellular homeostasis and disease states. Determining which protein is deacetylated by which HDAC isozyme is central to understanding biological regulation. To better identify HDAC substrates, I developed an assay to measure the acetate product formed by deacetylation catalyzed by metal-dependent HDACs for the evaluation of peptide substrates. Using this assay, I advanced a computational algorithm that predicts HDAC8 peptide substrates based on short range interactions. This algorithm accurately predicts the catalytic efficiency of 7-mer peptide substrates based on the sequence of the peptide. Using the deacetylation assay, I also demonstrated the reactivity of HDAC8 with peptide substrates derived from proteins with increased acetylation *in vivo* upon treatment with HDAC8 specific inhibitors. These experiments suggest that a subset of *in vivo* HDAC8 substrates can be predicted based on the six amino acids flanking the acetyl lysine. I have also utilized singly acetylated histone tetramers to establish that HDAC8 has enhanced activity in comparison to corresponding peptide substrates. Combined with the peptide work, these results suggest that molecular recognition by HDAC8 is determined by a combination of short and long-range interactions and acetyl lysine accessibility of *in vivo* substrates. Additionally, I have identified slow product dissociation as a novel regulatory method for HDAC activity. In the future we will expand these methods to identify substrates for other HDACs.

## Chapter 1

### HDAC8 substrates: Histones and beyond<sup>1,2,3</sup>

#### Overview

The lysine deacetylase family of enzymes (HDACs) was first demonstrated to catalyze deacetylation of acetyl lysine residues on histones. In subsequent years, HDACs have been shown to recognize a large pool of acetylated non-histone proteins as substrates. Recently, thousands of acetylated proteins have been discovered, yet in most cases, the HDAC that catalyzes deacetylation *in vivo* has not been identified. This gap has created the need for better *in vivo*, *in vitro*, and *in silico* approaches for determining HDAC substrates. While HDAC8 is the best kinetically and structurally characterized HDAC, few efficient substrates have yet been substantiated *in vivo*. In this review we delineate factors that may be important for determining HDAC8 substrate recognition and catalytic activity, including structure, complex formation, and post-translational modifications. This summary provides insight into the challenges of identifying *in vivo* substrates for HDAC8, and provides a good vantage point for understanding the variables important for predicting HDAC substrate recognition.

---

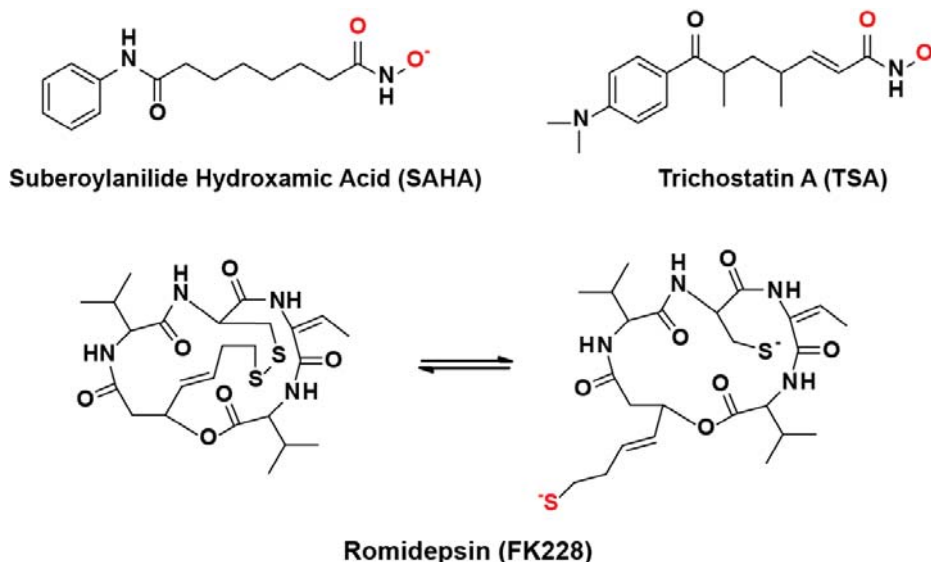
<sup>1</sup> Reproduced in part from Wolfson, N.A. Pitcairn, C.A. Fierke, C.A. 2013. HDAC8 substrates: Histones and beyond. *Biopolymers*. 2013 Feb;99(2):112-26. doi: 10.1002/bip.22135.

<sup>2</sup> Chapter was updated to include material published through 5/10/14.

<sup>3</sup> Noah Wolfson wrote the entire chapter except for the post-translational modifications of HDAC8 section which was written by Carol Ann Pitcairn.

## **HDAC introduction**

Acetylation of lysine side chains in proteins is a reversible post-translational modification that occurs in a wide range of organisms[1]. This modification affects the properties of proteins, including protein-protein association, protein-DNA interactions, and protein stability[2]. Initially, acetylation gained recognition as a post-translational modification to histones. Acetylation of histones can regulate the accessibility of DNA to cellular machinery and thus change the protein expression profiles of cells[3]. Due to the effect of acetylation on the proteome, it is not surprising that many diseases have been associated with the aberrant acetylation of histones[4]. In the last twelve years the paradigm for protein acetylation has changed drastically, moving from a histone centric model to a proteome centric model. This change in mindset has resulted from the identification of acetylated lysine side chains that affect the function of numerous non-histone proteins[5, 6]. Currently, over 3,600 acetylation sites have been discovered in mammalian proteins[7] and these protein are important in many cellular processes, including gluconeogenesis and DNA damage repair[5, 6]. Regulation of the acetylation state of proteins is important as aberrant acetylation of both histone and non-histone proteins can contribute to the development of many disease states[8-10]. As proof of this, two broad spectrum HDAC inhibitors (suberoylanilide hydroxamic acid (SAHA) and Romidepsin) have been approved by the FDA and are currently on the market for the treatment of T-cell Lymphomas (Figure 1.1)[11].



**Figure 1.1 Widely used HDAC inhibitors**

SAHA and Romidepsin have been approved by the FDA for use as second line treatments for T-cell lymphomas. TSA is an inhibitor that has been used widely in *in vitro* and *in vivo* studies but is not being tested in drug trials[12]. All three inhibitors are competitive with substrates by occupying the substrate binding channel and coordinating the active site metal ion. The atoms colored red interact with to active site metal ion[13-19].

HDAC isozymes can be grouped into four classes based on their phylogenetic similarity[20]. Class I (HDAC1, 2, 3, and 8), class IIa (HDAC4, 5, 7, and 9), class IIb (HDAC6 and 10), and class IV (HDAC11) enzymes catalyze deacetylation using a metal dependent mechanism[20, 21], while class III (Sirt1-7) enzymes use an NAD<sup>+</sup> cofactor to perform deacetylation[22, 23]. Due to the abundance and importance of HDAC substrates, one of the foremost questions in the field is the determination of the substrate specificity of HDACs. This area of research seeks to identify which of the 18 deacetylases catalyzes deacetylation of each of the >3,600 mammalian acetylation sites. Adding to the complexity of this problem is the possibility that cellular regulation may alter both the catalytic activity and the substrate specificity of HDACs. Illuminating the substrate selectivity and regulation of HDACs should shed light on the mechanism and treatment of acetylation-related diseases.

Mechanistically and structurally, HDAC8 is the best studied of the HDAC homologues. Furthermore, HDAC8 is proposed to recognize a number of non-histone substrates[24-26] and is

therefore a good model for developing techniques to unravel HDAC substrate specificity. In this chapter I discuss the current view of HDAC8 regulation, and compare HDAC8 to other promiscuous enzymes to identify factors that determine substrate specificity.

### **Known HDAC8 substrates**

HDAC8 was initially discovered in 2000 and was shown to catalyze *in vitro* deacetylation of a number of acetylated histone variants[27-29]. These substrates include full-length H2A/H2B, H3, and H4 histones, acetylated at non-specific lysines[27, 28]. Concurrent studies showed that peptide sequences corresponding to the H3 and H4 histone tails with an acetylated lysine at position fourteen (Kac14) and sixteen (Kac16), respectively, were also *in vitro* substrates[28-30]. In subsequent years, several studies have used the H4 histone tail sequence as a peptide template to investigate the amino acid sequence preference of HDAC8 (discussed below)[31-33]. Recently, HDAC8 was demonstrated to catalyze *in vitro* deacetylation of a peptide corresponding to the Kac20 site on the H4 histone tail. However, HDAC8-catalyzed deacetylation of the Kac20 peptide is much slower than deacetylation of Kac16 peptides[34], suggesting that another HDAC isozyme may catalyze this reaction *in vivo*. Further studies have shown that HDAC8 also catalyzes deacetylation full length histones[27-29]. Using western blots analysis, Van den Wyngaert *et al.* showed that cell extracts from HEK293 cells which over express HDAC8 have significantly lower levels of acetylated H4 and modestly reduced levels of acetylated H3[29]. Similarly, Lee *et al.* showed that HDAC8 catalyzes the deacetylate acetylated histone H3 and H4 *in vitro*[35]. In contrast, Olson *et al.* show that treatment of HeLa cells with a cell permeable HDAC6/8 specific inhibitor (PCI-34051)[36] does not significantly alter the levels of H3 acetylation compared to treatment with the pan-HDAC inhibitor SAHA[37], yet TSA has been shown to have an effect on H4 acetylation levels[38]. Taken together, these data indicate that HDAC8 catalyzes the deacetylation the H3 and H4 histones *in vitro*, while HDAC8 catalyzed deacetylation of *in vivo* substrates remains controversial. The discrepancies between the *in vivo* and *in vitro* data may be due to a number of factors including: the ability of other HDACs to catalyze the deacetylation of H3; cofactors and post translational modification alter the reactivity of HDAC8 with histone H3 *in vivo*; or that other factors, including cell cycle



progression, mask HDAC8's effects on H3. Additionally, there are data suggesting that HDAC8 may affect the acetylation of histone H3 in specific gene regions[39] and thus the HDAC8 modulation of H3 acetylation (either directly or indirectly) may be masked by the noise of other regions on chromatin. As a result, the role of HDAC8 in catalyzing deacetylation of specific sites in histones *in vivo* remains unclear.

Shortly after HDAC8 was identified, the first non-histone acetylated proteins were reported[40, 41] which inspired researchers to hunt for other possible HDAC substrates. The search for new HDAC8 substrates was further spurred by the finding that this enzyme is present in the cytoplasm of smooth muscle cells[42, 43], causing evaluation of non-nuclear substrates. In fact, HDAC8 catalyzes deacetylation of a peptide corresponding to the C-terminal end of the p53 transcription factor (Figure 1.2A) faster than the Kac16 H4 histone peptide[Biomol unpublished]. HDAC8 catalyzes deacetylation of coumarin derivatives of the acetylated p53 and H4 peptides with  $k_{cat}/K_M$  values of  $7,500 \text{ M}^{-1}\text{s}^{-1}$ [26] and  $2,800 \text{ M}^{-1}\text{s}^{-1}$ , respectively[44]. Since the  $k_{cat}/K_M$  parameter reflects the relative reactivity of an enzyme with different substrates[45], these values suggest that HDAC8 has a modest preference for catalyzing deacetylation of p53 over the H4 histone. It is important to note that these  $k_{cat}/K_M$  values for HDAC8 were measured using the commercially available Fluor-de-lys assay (Biomol) (Figure 1.2B). This assay uses peptide substrates containing a methylcoumarin fluorophore conjugated to the C-terminal side of the acetyl lysine residue. After deacetylation, digestion by trypsin cleaves the coumarin fluorophore, causing an increase in fluorescence at 460 nm and deacetylation is measured from this increase in the fluorescence signal[46] (Figure 1.2B). While this assay has been a valuable tool for studying histone deacetylases, the methylcoumarin fluorophore increases the reactivity with HDAC8[47]. Therefore, deacetylation of the nonlabeled acetylated p53 and H4 histone peptides catalyzed by HDAC8 may be slower than reported using this assay. Furthermore, the coumarin substrates may not reliably reflect HDAC substrate specificity in the context of full-length proteins.

A

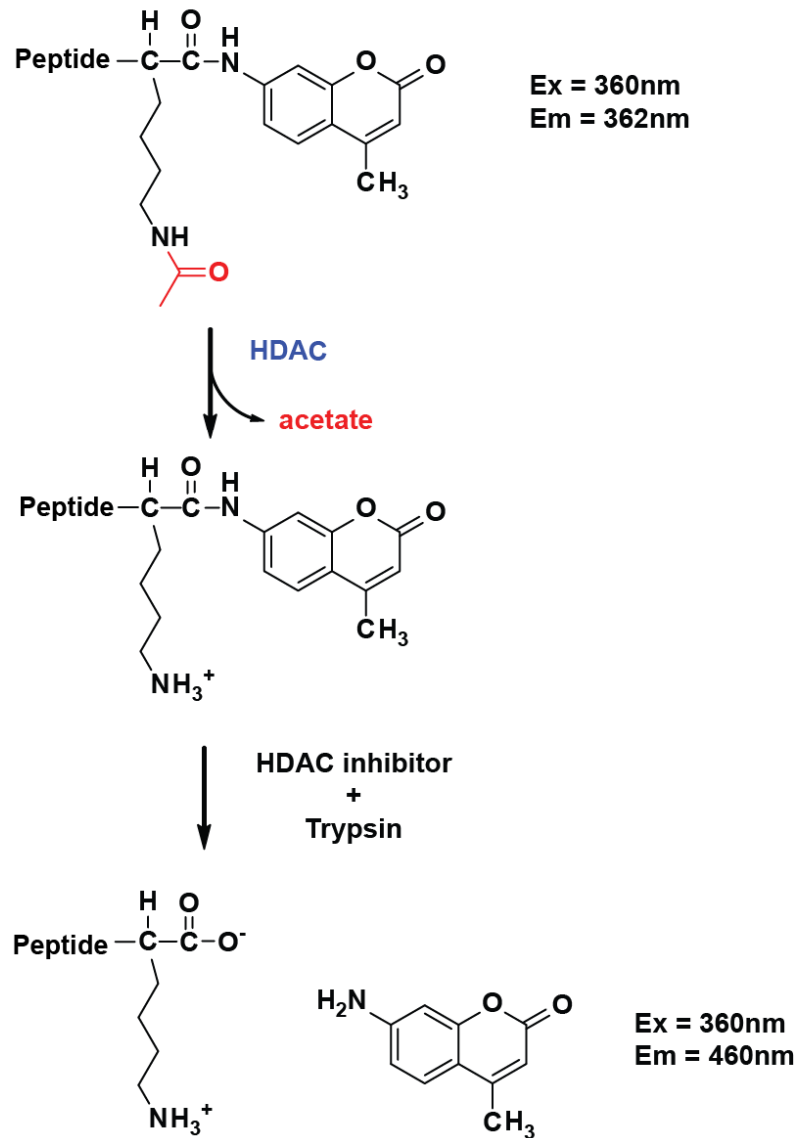
**Peptide Representing p53**

Ac-Arg-His-Lys(ac)-Lys(ac)-methylcoumarin

**Peptide Representing H4**

Ac-Lys-Gly-Gly-Ala-Lys(ac)-methylcoumarin

B



**Figure 1.2 The Fluor de Lys assay[Biomol][46]**

A) The sequence of two HDAC8 substrates used for the Fluor de Lys assay. B) Schematic of the Fluor de Lys assay, including the wavelengths used to measure the methylcoumarin fluorophores.

The steady state kinetic parameters for catalysis of the deacetylation of peptides can provide insight into both the kinetic mechanism and the *in vivo* reactivity of these substrates. HDAC8-catalyzed deacetylation of the p53 and H4 coumarin peptides have a low value of  $k_{cat}/K_M$  ( $10^3 - 10^4 \text{ M}^{-1}\text{s}^{-1}$ ) in comparison to enzymes that function near diffusion-controlled limits ( $10^6 - 10^8 \text{ M}^{-1}\text{s}^{-1}$ ) and a high value for  $K_M$  (320  $\mu\text{M}$ , H4 peptide)[44] compared to other HDAC isozymes ( $\sim 30 \mu\text{M}$ )[48]. These data suggest a simple Michealis-Menten kinetic model whereby substrate binding and dissociation is rapid, and is followed by rate-limiting deacetylation. This conclusion is bolstered by the observed enhancement of the  $k_{cat}$  value for deacetylation of peptides labeled with a more reactive trifluoroacetyl group[30, 49]. Therefore, substrate specificity is determined by both the affinity of HDAC8 for a peptide substrate and the reactivity of the enzyme-substrate complex.<sup>4</sup> Assuming that the kinetic constants for deacetylation of these peptides mimic the full-length proteins, the low  $k_{cat}/K_M$  and high  $K_M$  values for the H4 and p53 peptides compared to reactivity with other isozymes[26, 44, 48] suggest that HDAC8 may not catalyze deacetylation of these sites *in vivo*. However, it is possible that natural, full length, substrates may be better optimized for efficient deacetylation to allow for regulation of these posttranslational modifications. While knockdown data suggest that HDAC8 levels affects acetylation levels of K382 of p53 *in vivo*[50], other data implicating HDAC8-catalyzed deacetylation of H4 and p53 *in vivo* is sparse. In addition to these proposed substrates, *in vitro* kinetic studies combined with cellular assays have yielded several promising candidates for *in vivo* HDAC8 substrates (discussed further below).

There are a number of factors that must be taken into account when parsing whether substrates are acted upon by a given enzyme *in vivo*. HDAC selectivity is minimally described by the relative values of  $k_{cat}/K_M$  for deacetylation, the relative concentrations of the HDAC isozymes, and the concentrations of competing substrates. Relative  $k_{cat}/K_M$  values indicate the substrate preference of an enzyme when discriminating amongst multiple substrates[45]. The majority of enzymes have  $k_{cat}/K_M$  values of  $10^5$ - $10^6 \text{ M}^{-1}\text{s}^{-1}$ [51]. These values are generally slower than the diffusion controlled rate constants for substrate binding, which can be as high as  $10^7$ - $10^8 \text{ M}^{-1}\text{s}^{-1}$ [45]. Consistent with this, the  $k_{cat}/K_M$  values for the HDAC8 homolog HDAC1 and, the

---

<sup>4</sup>  $K_M$  and  $K_D$  are roughly equivalent as data suggests that the rate limiting step for HDAC8 catalysis is chemistry (see Chapter 5 for a more through explanation).

homologous enzyme, arginase I are on the order of  $10^5 \text{ M}^{-1}\text{s}^{-1}$ [48, 52], suggesting that similar values should be achievable for efficient HDAC8 substrates. One caveat to making conclusions from kinetic parameters measured in *in vitro* experiments is that some enzymes require an activator for optimal activity. Since many HDAC isozymes associate with large protein complexes *in vivo*, it is possible that other proteins in the complex could activate the catalytic activity or enhance the substrate affinity to increase the value of  $k_{\text{cat}}/K_{\text{M}}$  in the cell.

### **Candidate non-histone HDAC8 Substrates**

One promising HDAC8 substrate is the Estrogen-Related Receptor  $\alpha$  (ERR $\alpha$ ). This orphan receptor is expressed in a number of organs, including the heart, kidney, and muscle, where it controls processes that are essential for maintaining energy homeostasis[53]. ERR $\alpha$  can be acetylated at four lysines, where these post-translational modifications inhibit DNA binding[25]. A role for HDAC8 in catalyzing the deacetylation of ERR $\alpha$  was suggested by the demonstration that the acetylation state of ERR $\alpha$  was altered by simultaneous incubation with HDAC8, the histone acetyltransferase PCAF, and  $^{14}\text{C}$ -acetyl-CoA[25]. Furthermore, incubation of purified acetylated-ERR $\alpha$  with HDAC8 enhances the affinity of ERR $\alpha$  for DNA, which is consistent with HDAC8-catalyzed deacetylation of ERR $\alpha$ . One caveat to these experiments is that this assay included metal chelators and low salt, conditions where HDAC8 has limited catalytic activity[26, 54]. An alternative explanation of these data is that HDAC8 binds to ERR $\alpha$  to increase the DNA affinity and decrease acetylation catalyzed by PCAF. However, addition of the non-homologous deacetylase, Sirt1, to these *in vitro* assays also decreases acetylation of ERR $\alpha$ , suggesting that both enzymes recognize ERR $\alpha$  as a deacetylase substrate. Finally, RNAi-dependent decreases in cellular HDAC8 or Sirt1 levels are accompanied by increased ERR $\alpha$  acetylation *in vivo*[25]. Taken together, these results suggest that HDAC8 catalyzes deacetylation of ERR $\alpha$  *in vivo*. Consistent with this, the acetylation site (K129ac) in ERR $\alpha$  has Arg in the -1 position (the amino acid on the N-terminal side of the acetyl lysine), and RKac motifs have been demonstrated to be favorable for HDAC8 catalysis[47]. Additional analysis, such as directly measuring ERR $\alpha$  acetylation patterns using mass spectrometry in the presence and absence of HDAC inhibitors would further validate ERR $\alpha$  as an *in vivo* substrate of HDAC8.

Another proposed HDAC8 substrate is the aberrant *inv(16)* fusion protein found in a significant portion of patients with acute myeloid leukemia[55]. This fusion protein combines the N-terminus of the transcription factor domain core binding factor  $\beta$ , with the C-terminus of the smooth muscle myosin heavy chain[56]. In COS7 cells, co-immunoprecipitation experiments demonstrated that overexpressed HDAC8 associates with *inv(16)*[24]. Furthermore, HDAC8 co-localizes and immunoprecipitates with smooth muscle myosin heavy chain[57] suggesting that HDAC8 may interact with this domain within the *inv(16)* fusion protein. Other HDAC isozymes do not immunoprecipitate with *inv(16)* under similar conditions, which suggests that HDAC8 may be the main HDAC that interacts with *inv(16)* *in vivo*. The addition of the HDAC inhibitor TSA inhibits the transcriptional repression activity of *inv(16)*[24], suggesting that HDAC8 activity is important for *inv(16)* regulation. An alternative explanation of these data is that *inv(16)* is a binding partner with HDAC8 rather than a substrate, as HDAC inhibitors have been shown to disrupt the association of HDACs with non-substrate binding partners[58]. The sequence of the acetylation site in the core binding factor  $\beta$  is RSKacFE[5]. Peptide library studies have demonstrated that Phe in the +1 position is favorable for HDAC8 catalysis[32, 47] although Ser at the -1 position attenuates reactivity[47]. Furthermore a computational algorithm for determining HDAC8 substrates ranks this peptides as a substrate (Chapter 3). While the core binding factor  $\beta$  is acetylated *in vivo*[5], there is not yet direct evidence that *inv(16)* is acetylated[59]. Taken together, these data indicate that *inv(16)* is either an HDAC8 substrate or forms a functionally important complex with HDAC8.

A third potential *in vivo* HDAC8 substrate is the transcription factor CREB. Acetylation at three CREB sites (Lys91, Lys96, and Lys136) helps to activate this protein[60]. HDAC8 and CREB overexpressed in HEK293 cells co-immunoprecipitate, demonstrating that these two proteins associate. When HDAC8 is overexpressed in cells, phosphorylation of CREB decreases, which in turn inhibits CREB transcriptional activation[61]. Likewise, treatment of cells with the HDAC inhibitor TSA increases CREB phosphorylation levels[62] suggesting that HDAC8 activity is important for CREB phosphorylation. However, the addition of a broad range HDAC inhibitor (such as TSA) decreases the activity of all metal-dependent HDACs. As HDAC overexpression can affect a number of targets within the cell, this inhibition may indirectly affect CREB phosphorylation. Furthermore, pulldown experiments demonstrate that CREB can interact with a

number of HDAC isozymes[61], complicating identification of CREB as an HDAC8 substrate *in vivo*. Due to the high amino acid identity between class I HDACs (>30%)[63], overexpression and pulldown experiments may not yield results that are representative of *in vivo* situations. Therefore, these experiments suggest, but do not confirm, a direct connection between HDAC8 deacetylase activity, the phosphorylation status of CREB, and the regulation of CREB activation. Alternatively, HDACs may function as protein scaffolds to mediate the inhibitory interaction between CREB and PP1 phosphatase[61, 64-66], leading to a decrease in CREB phosphorylation and activity.

The cohesin subunit SMC3 is a putative *in vivo* substrate that has gained acceptance by the HDAC field. The first evidence for this was the demonstration that 5-30% of Cornelia de Lange syndrome (CdLS) patients suffer from mutations in HDAC8[67, 68], while the vast majority of these patients suffer from mutations to the cohesin complex or associated proteins. As a result, the link between HDAC8 and cohesin was investigated. SMC3 acetylation levels were shown to rise upon HDAC8 siRNA knockout or inhibition by an HDAC8 specific inhibitor[69]. This suggests that HDAC8 either deacetylates SMC3 directly or regulates another acetyl transferase or deacetylase that acts on SMC3. siRNA knockout of HDAC8 effects cohesin localization within the chromosomes, showing that HDAC8 affects cohesin function. HDAC8 mutants found to cause CdLS phenotypes in humans were shown to reduce HDAC8 activity and cause gene expression phenotypes similar to those seen with other CdLS mutations. HDAC8 mutants which cause CdLS reduce HDAC8 activity by varying degrees; some mutations fully kill HDAC8 activity while others lessen it by two fold[68, 69]. The location of the various mutants within the HDAC8 structure suggest variable effects on activity including: altering substrate specificity; altering protein stability; and altering HDAC8 binding partner affinity.

The current cellular methods for identifying substrates of HDAC isozymes *in vivo* have limitations. Since HDAC selectivity depends on the relative concentrations of the HDAC isozymes and the concentrations of all of the acetylated lysine substrates, overexpression of HDAC and/or HDAC substrates can alter the normal pattern of deacetylase activity. Therefore, experiments using overexpressed proteins can suggest that a particular interaction occurs *in vivo*, but does not prove that this contact occurs under physiological conditions. Native pulldown experiments, which should be more representative of physiological conditions, have thus far not

been successfully used to confirm the identity of HDAC8 substrates. It is possible that the HDAC-substrate interactions may be transient and/or weak and thus are not maintained through the multiple washes in pulldown experiments. Therefore alternate techniques, such as crosslinking, may be necessary to increase the lifetime of an HDAC-substrate complex to allow detection. Additionally, observation of enhanced acetylation after deletion or knockdown of a given isozyme does not prove that an HDAC isozyme directly catalyzes deacetylation of that site. Therefore alternative methodologies need to be explored to enhance the identification of additional HDAC8 substrates. One novel technique used for the identification of HDAC8 substrates is mass spectrometry tracking acetylation with HDAC specific inhibitors (Chapter 4). This methodology compares the acetylation state of proteins after a cell has been treated with an HDAC specific inhibitor and a control. The differential acetylation state of identified proteins suggest that they are substrates, though this method does not differentiate between substrates and proteins affected downstream by HDAC deacetylation.

### **HDAC8 complex formation**

Recombinantly purified HDAC8 catalyzes deacetylation and displays peptide substrate selectivity in the absence of additional protein cofactors[16, 17, 26-28, 30-33, 47, 48, 54, 70], suggesting that HDAC8 can catalyze deacetylation *in vivo* in the absence of a protein complex. In contrast, the other class I HDACs, HDAC1, 2, and 3, are observed in complexes in the cell and their substrate specificity largely depends on the combination of proteins incorporated into their complexes[71]. HDAC1 and 2 associate with Sin3 scaffolded complexes which serve a range of functions within the cell. The substrate specificity and function of these HDAC isozymes can change by altering the protein composition of the complex[72]. Although HDAC8 is phylogenetically most similar to the other class I HDACs, divergent evolution[20] may have altered how HDAC8 interacts with cofactors, possibly allowing this isozyme to function independent of other proteins. However, HDAC8 does associate with other proteins, and these interactions may affect the biological function and selectivity of this enzyme.

Distinguishing between HDAC8 substrates and binding partners in the cell is currently difficult, as discussed in the previous sections. For example, previous experiments have provided evidence

that the HDAC1/HDAC2 complex associates with both the PP1 phosphatase and CREB, leading to decreased CREB phosphorylation[66]. Because an inactive HDAC1 mutant still affects CREB activity, the function of the HDAC1/HDAC2 complex was proposed to co-localize PP1 phosphatase and CREB. However, it is possible that HDAC2 catalyzes deacetylation of CREB under these conditions[66]. Similarly, both PP1 and CREB co-immunoprecipitate with HDAC8, and HDAC8 overexpression decreases CREB activity. These data are consistent with HDAC8 either acting as a scaffold to enhance the interaction between PP1 phosphatase and CREB or catalyzing deacetylation of CREB.

In addition to CREB, HDAC8 likely affects the activity of other transcription factors either through their deacetylation or through complex formation. Levels of the transcription factor DEC1, were shown by DNA-CHIP analysis to correlate with the co-localization of HDAC8 at specific promoter sites[73]. Co-immunoprecipitations of HDAC8 and DEC1 validate their association, though it is unclear whether HDCA8 catalyzes the deacetylation of DEC1.

Furthermore, HDAC8 expression can change DEC1's expression profile, where over-expression of or siRNA knockdown of HDAC8 correlates with the up- or down-regulation (respectively) of the DEC1 regulated expression levels of the TAp73 transcription factor. This result led to the suggestion that DEC1 recruits HDAC8 to the TAP73 promoter. Additionally, HDAC8 has been shown to affect the mRNA and expression levels of a number of other genes such as SOCs1, SOCS3[74, 75], and p53[76]. Based on these results, HDAC8 may act in a complex which is shuttled to DNA regions where it can act on nucleosomes and nucleosome-associated protein to alter gene expression. Alternatively, HDAC8 may stably associate with some substrates and affect their acetylation, thus changing their activity. Either way, it appears that HDAC8 has the ability to alter a subset gene expression.

HDAC8 also co-localizes with  $\alpha$ -actin, as indicated by immunofluorescence staining[42]. This interaction was confirmed by pulldown experiments using human smooth muscle cells, demonstrating an endogenous association between  $\alpha$ -actin and HDAC8[43, 57]. The function of this interaction was partially elucidated by demonstrating that siRNA knockdown of HDAC8 in human smooth muscle cells decreased the ability of cells to contract, when exposed to a collagen lattice. Furthermore, the siRNA-treated smooth muscle culture cells were smaller and unable to spread. These changes in cell morphology occurred without detectable changes to  $\alpha$ -actin



acetylation[43], suggesting that HDAC8 may act as part of a complex which modulates the cell cytoskeleton. Furthermore, pulldown experiments demonstrate that HDAC8 associates with the proteins Hsp20, myosin heavy chain, and cofilin[57] all of which can potentially affect actin dynamics[77, 78]. It is currently unclear whether Hsp20 or cofilin are acetylated and substrates for HDAC8. However, immunoprecipitation experiments suggest that HDAC8 associates better with the non-acetylated form of myosin heavy chain, suggesting that this protein is not an HDAC8 substrate[57]. Furthermore,  $\alpha$ -tubulin (another cytoskeletal protein) acetylation levels are altered upon either addition of an HDAC8 specific inhibitor[38] or HDAC8 siRNA knockdown[79] which in turn destabilizes the microtubule system. Because HDAC8 enhances cell contractility and associates with many proteins important for actin function, it is likely that HDAC8 is a component of a complex that modulates actin and microtubule dynamics. The relationship between microtubule and cohesion activity[80-82] may provide an alternative explanation for HDAC8's effects on microtubule dynamics.

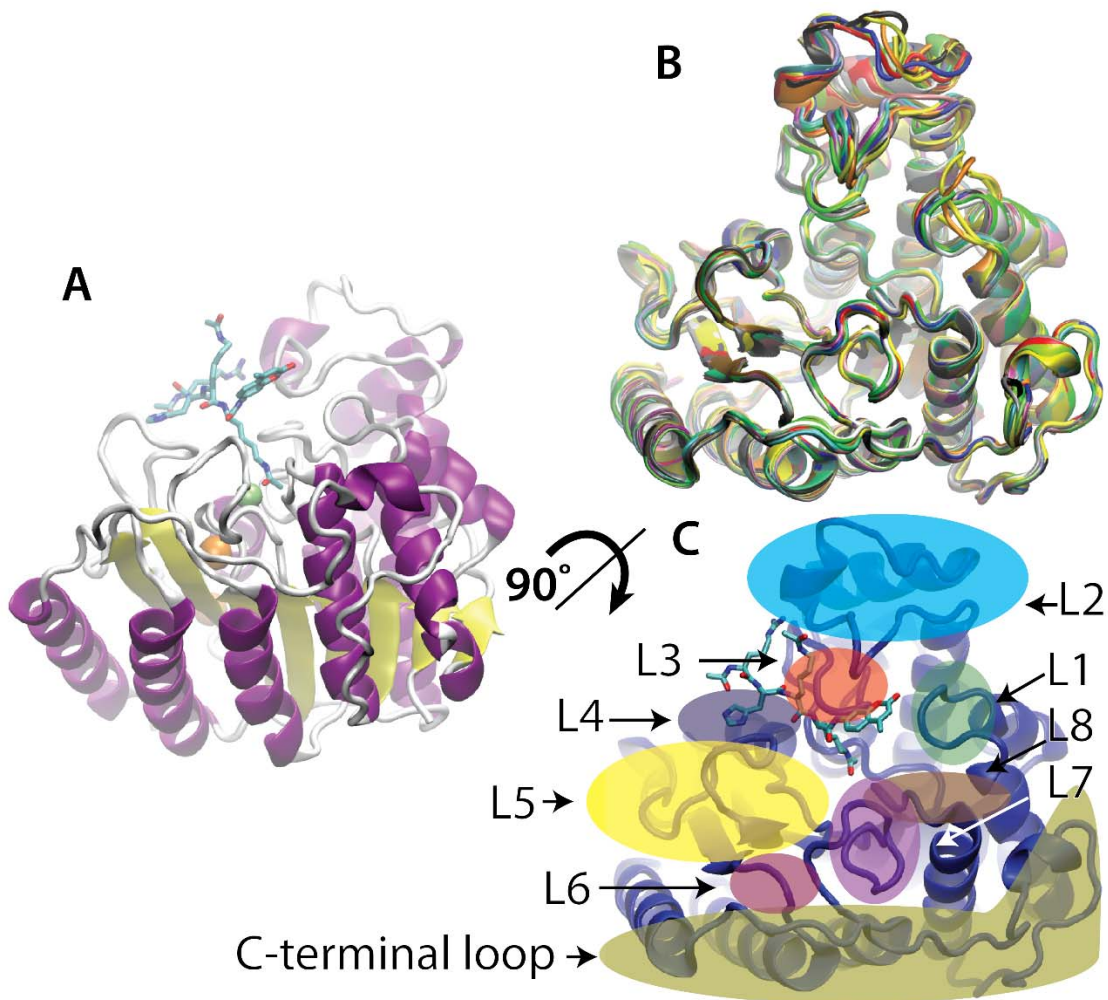
Additional potential HDAC8 interaction partners have been identified using a bacterial two-hybrid system[83]. Two of the fifteen identified binding partners have been examined in detail: the human Ever-Shorter Telomeres 1B (hEST1B) protein that activates telomerase activity and HOP1, an adaptor protein linking Hsp70 and Hsp90. The two-hybrid results were confirmed using co-immunoprecipitation of overexpressed hEST1B and HDAC8 in HeLa cells. HDAC8 knockdowns led to decreased telomerase activity through diminished levels of hEST1B. As HDAC8 activity does not affect the promoter region regulating hEST1B, the hEST1B level is likely not regulated by alteration in transcription. However, hEST1B levels are increased by addition of a proteasome-dependent pathway inhibitor or decreased by overexpression of ubiquitin and rescued by phosphorylated HDAC8. These results argue that phosphorylated HDAC8 protects hEST1B from polyubiquitination and subsequent degradation by the proteasome. Since phosphorylation decreases HDAC8 activity this protective effect could be due to enhanced acetylation of hEST1B. The protective effects of phosphorylated HDAC8 on hEST1B levels are independent of deacetylase activity, remaining in the presence of the catalytically inactive His143Ala-HDAC8 mutant, or after exposure of cells to TSA. Therefore, HDAC8 interacts with hEST1B but deacetylation is not required for the functional effect. To further explore the interaction between HDAC8 and HOP1 indicated by the two-hybrid

experiment, the association of HDAC8 with known HOP1 binding partners was investigated. Pulldown experiments demonstrated that endogenous Hsp70 and Hsp90 co-immunoprecipitate with overexpressed HDAC8[83]. This result suggests that HDAC8, HOP1, Hsp70, and Hsp90 form a complex. One proposed mechanism for the effect of HDAC8 on telomerase activity suggests that the Hsp70-HDAC8 complex protects hEST1B from ubiquitination catalyzed by the E3 ubiquitin ligase CHIP[83]. This in turn raises the levels of hEST1B and activates telomerase. Interestingly, interaction of HDAC8 with the Hsp proteins may help to elucidate the effect of HDAC8 on  $\alpha$ -actin, since Hsp90 has been proposed to modulate  $\alpha$ -actin dynamics[84, 85]. Thus it is possible that the HDAC8-HOP1-Hsp90 complex may regulate  $\alpha$ -actin function.

A similar study to the two-hybrid system (above) identified proteins that associate with an HDAC8-GFP fusion using mass spectrometry[86]. These data solidify several previously suggested HDAC8 interaction partners such as SMC1A and SMC3, while identifying a number of additional HDAC8 interaction partners/substrates. Some of these partners further elucidate proposed HDAC8 functions, while others hint at HDAC8 functions which have yet to be identified. Joshi *et al.* identified a number of cohesion associated proteins (STAG2, SMC1A, SMC3) and microtubule associated proteins (TPM3, TPM4, ADD3, SVIL) as well as transcription factors (ZSCAN2, MDM1, ZSCAN2). These proteins solidify HDAC8's role in regulating cell motility and transcription. Other associated protein suggest a role for HDCA8 in additional cellular functions. For example associations of HDAC8 with mitochondrial protein CHCHD3 and the calcium dependent membrane protein CPNE3, suggest a role for HDAC8 in calcium mediated apoptosis[36] (as previously proposed). HDAC8 shares interaction partners with other HDACs. For example both HDCA8 and HDAC6 interact with microtubules and protein transport proteins, consistent with pervious data[42, 43, 87]. However Joshi *et al.* found limited interactions between the binding partners of HDAC8 and another isozymes. While this phenomenon could be an artifact, it could indicate that HDAC8 specializes in its functions and as a result may also have a relatively exclusive set of substrates. Finally, experimental variation of binding partners between trials could indicate that regulation affects HDAC8 activity and function. This phenomenon could explain why Joshi *et al.* failed to identify many of the HDAC8 binding partners found in other studies. Alternatively, the GFP tag on HDAC8 may have prevented number of complexes from forming.

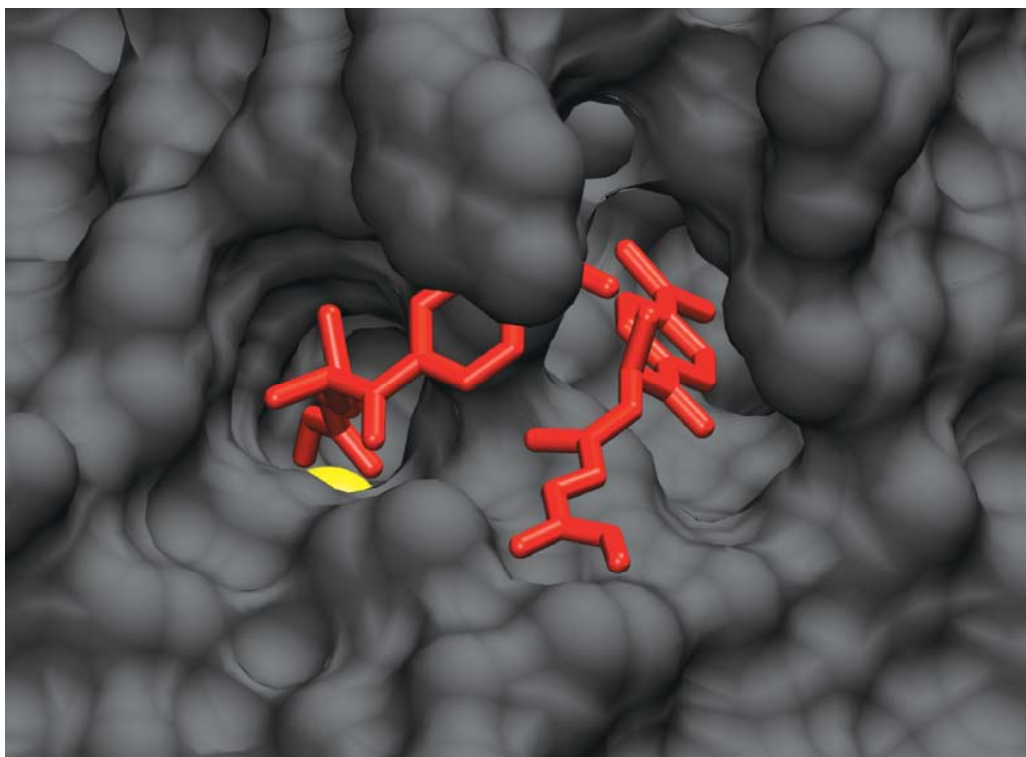
## Enzyme structure affects substrate specificity

The structure of HDAC8 yields clues about molecular recognition relevant to substrate selectivity. HDAC8 is the second smallest metal-dependent HDAC at ~42 kDa, containing little more than the catalytic domain[20, 27-29]. This HDAC folds as a single  $\alpha/\beta$  domain with a core eight-stranded  $\beta$ -sheet surrounded by eleven  $\alpha$ -helices (Figure 1.3A). The substrate binding surface, composed of nine loops and an 11 Å tunnel leading to the active site, is proposed to be conformationally flexible based on the poor occupancy and varying positions of the loop residues in crystal structures[13-19] (Figure 1.3B). Furthermore, one crystal structure illuminates a bound TSA molecule interacting with residues in the hydrophobic core of HDAC8[14] (Figure 1.4). While this may simply be an artifact, the alternative binding mode suggests that the surface of the protein can change conformation enough to allow hydrophobic molecules to intercalate between these loops and interact with the interior of the protein. Loops are a common structure in promiscuous enzymes[88] and examples of proteins, such as chymotrypsin[89] and carboxypeptidase A[90], that use loops to bind a range of substrates are abundant in nature. These loops create a number of different conformations that bind ligands through a combination of induced fit and select fit mechanisms[45, 91]. The varied conformations and motifs provide a palette of binding sites to accommodate a multiplicity of substrates, and may account for the complex dynamics observed and simulated for HDAC8 inhibitors[92-94]. Furthermore, long range allosteric movements propagated through the loops may affect the active site and surrounding areas, potentially altering substrate preferences.



**Figure 1.3 HDAC8 structures**

A) Side view of HDAC8 with bound peptide substrate PDBID: 2v5w[13]. Helices are purple, sheets are yellow, turns are white, the monovalent cations are orange, and the active site metal is colored green. The Fluor de Lys substrate representing the p53 sequence is colored cyan for carbon, red for oxygen, and blue for nitrogen. B) Front view of an overlay of the 21 HDAC8 crystal structures in the PDB: PDBID: 2v5x, 2v5w, 1t69, 1t64, 1vkg, 1t67, 1w22, 3sfh, 3sff, 3mz3, 3ezt, 3fo6, 3mz4, 3mz6, 3mz7, 3ew8, 3ezp, 3f07, 3f0r, 3ewf, and 3rqd[13-19]. Structural variations are especially apparent in the L1, L2, and C-terminal loops. C) A map of the crystal structure of HDAC8 outlining the loop regions.



**Figure 1.4 HDAC8 with two bound TSA molecules**

PDBID: 1t64[14]. In this crystal structure, one molecule of TSA binds to the active site tunnel to coordinate the divalent metal ion (colored yellow) while a second TSA molecule binds nearby in between the L1, L2, and L3 loops.

In fourteen of the twenty-one HDAC8 crystal structures, the enzyme crystallizes as a dimer along the substrate binding interface[13-18]. As HDAC8 is a monomer in solution[19], the dimer interface may provide insight into long range interactions between HDAC8 and its *in vivo* substrates. To date, substrate specificity has mainly been evaluated using peptide substrates, therefore only short range interactions have emerged as HDAC8 substrate binding motifs[31-33, 47]. Based on the crystal structure of bound peptides[13, 16] and biochemical measurements, these interactions include ring stacking, hydrogen bonding, salt bridges, and electrostatic interactions. Ring stacking between Tyr100 and the methylcoumarin of the Fluor-de-lys peptides is observed in the crystal structure[13, 16]. Similarly, ring stacking between aromatic amino acids in the +1 position and Tyr100 may be important for substrate recognition[32, 47]. Additionally, hydrogen bonding between the backbone amides of the substrate and the Asp101

side chain oxygens are important for molecular recognition[16]. Salt bridges between positively charged arginines in the substrate and negatively charged carboxylate side chain oxygens, and general hydrophobic interactions can be seen in the peptide-enzyme interface[13, 16]. Because of the limited number of interactions, the binding affinity may be dominated by a few strong contacts, as observed for the interaction between Tyr100 of HDAC8 and the methylcoumarin moiety of short Fluor-de-lys peptides[47]. This pi-pi interaction (~2 kcal/mol)[95, 96] is of comparable energy with other HDAC8-peptide contacts. In contrast, binding a protein substrate could involve many more contacts, including multiple hydrogen bonds (0.5-1.5 kcal/mol), hydrophobic (~1 kcal/mol), electrostatic (<1 kcal/mol)[45], and solvent exposed salt bridge (~1-3 kcal/mol)[97] interactions. Therefore, the binding affinity could depend on a large number of interactions that together create a promiscuous substrate binding profile. Determinants of substrate specificity are still being evaluated for HDACs and further identification of binding motifs will be beneficial for understanding the biology of these enzymes.

When the structure of HDAC8 is compared to that of the homologous polyamine deacetylase, APAH[98], striking differences in loop size and structure can be observed. These differences in the loops may be important for substrate binding as APAH catalyzes deacetylation of small molecules, including acetylated spermidine, putrescine, and spermine, while HDAC8 deacetylates macromolecules. In APAH, the L1 and L2 loops are much larger and contain many more hydrophobic residues than the corresponding HDAC8 loops (Figure 1.5A,B), while the C-terminal loop and helix in HDAC8 are absent in APAH. Similarly, a comparison of the L1, L2, and C-terminal loops of different HDACs reveals interesting variations. The L1 and L2 loops of HDAC2[99], 4[100], 7[101], and 8[13] are more divergent in size, structure, and number of charged residues than other loops within these HDACs (Figure 1.5A,B). For instance, the size and number of charges within the L1 and L2 loops change two-fold between HDAC8 and HDAC4. Comparison of all of the HDAC8 crystal structures illustrates that the L1 and L2 loops have the most structural variability of the loops in the proposed substrate binding surface, suggestive of a role in ligand binding. Additionally, the L2 loop interacts with inhibitors, suggesting that it may be important for molecular recognition of substrates[16]. The L3 loop, which lies below the L2 loop and flanks the active site, also varies greatly in the number of charges in the loop among HDACs 2, 4, 7, and 8, consistent with a role in substrate or binding

partner selectivity. The C- and N- terminal portions of the HDACs, which lie on the outer edge of the substrate binding surface, may also interact with ligands. In the HDAC crystal structures, the C-terminal loops vary in position, charge, and size and may be responsible for long distance interactions between HDACs and their substrates, or used for recognition of binding partners.







### Figure 1.5 Structural variation within the structurally characterized HDACs

A) Aligned sequences of the published HDAC crystal structures[13, 98-101] using COBALT[102]. Highlighted in yellow are the residues that comprise the loop regions and amino acids that comprise the putative substrate binding region. The positively charged residues are red and the negatively charged residues are green. B) Surface visualizations of the crystal structures for HDAC2 (PDBID: 3max[99]), HDAC4 (PDBID: 2vqm[100]), HDAC7 (PDBID: 3c0y[101]), HDAC8 (PDBID: 2v5w[13]), and APAH (PDBID: 3q9b[98]). Superimposed into each structure is the Fluor de Lys substrate (green) from the HDAC8 structure. In red are the positively charged residues Arg and Lys, and in blue are the negatively charged residues Asp and Glu.

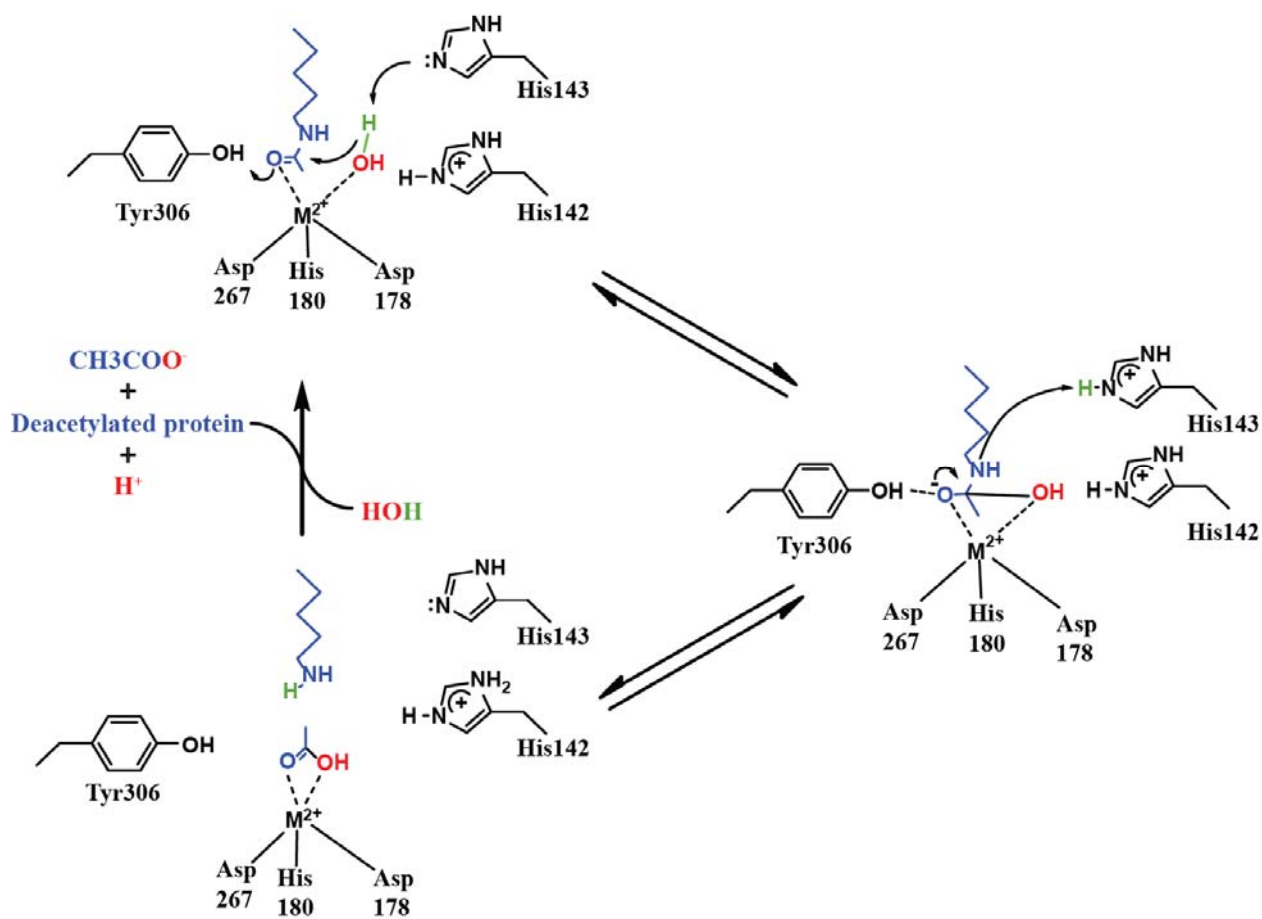
Along with structural studies, peptide substrates have been useful for evaluating substrate motifs recognized by HDAC8. Reister and colleagues measured the reactivity of HDAC8 with a peptide library of the sequence Ac-X-Z-Kac-methylcoumarin, where X and Z were all amino acids except for cysteine[31]. This work indicated that HDAC8 favors Pro, Met, Ala, Lys, Arg, Gln, Asp, Phe, and Ser at the -2 position and aromatic (Phe, Trp, and Tyr) and hydrophobic (Ile, Met, and Val) amino acids at the -1 position. However, the activity of HDAC8 in these assays was low, possibly due to the inclusion of the metal chelator EDTA in the assay. The Mrksich group developed a mass spectrometric assay to profile the local substrate specificities of HDACs[47]. The reactivity of HDAC8 with a peptide array of the sequence, Ac-G-X-Kac-Z-G-C-NH<sub>2</sub> where X and Z were any amino acid other than cysteine, showed that the most efficient substrate contains Arg and Phe at the X and Z positions, respectively[47]. However, HDAC8 also catalyzes deacetylation of peptides containing the sequence X=Arg/Z= variable and X = variable/Z = Phe. HDAC8 selectivity was further screened using a peptide library with the following sequence: Ac-G-R-Kac-**X-Z**-C-NH<sub>2</sub>[32]. These data demonstrated a preference for Arg or Phe at the X position. Furthermore, when X is Phe the identity of the Z position has only a modest effect on activity. These results suggest that specific positions and combinations of amino acids contribute significantly to the substrate recognition of small peptides, while other positions fine tune recognition. The Mrksich group also demonstrated that an RHR motif added to the C-terminus of peptide substrates of varying lengths enhances reactivity, demonstrating that distal sequences can modulate HDAC8 substrate selectivity[33]. Interestingly, the sequences RHRK and RHKK are found in the H4 histone tail and in p53, respectively, and hint that distal sequences may enhance the reactivity of HDAC8 with these substrates in cells.

Finally, the structure of the active site may also play a role in HDAC substrate specificity. HDAC2 and 8 have well defined 11Å channels leading to their active sites that easily accommodate an acetyl lysine side chain, however, this tunnel is lacking in HDAC4 and 7[13, 99-101] where only half of the channel is apparent. This modification in active site structure could suggest that HDAC4 and 7 catalyze deacetylation of alternate substrates, as proposed by Lombardi et. al.[98]. Alternatively, these isozymes might need substrates or binding partners that complement the active site to stabilize the binding of the acetyl lysine moiety.

### **Catalytic Mechanism and Regulation of HDAC8 activity**

The active site of HDAC8 contains a divalent metal ion coordinated to two aspartate and one histidine side chain (Asp178, Asp267, and His180) and one or two water molecules. Additionally, a conserved tyrosine (Tyr306) and a pair of conserved histidine/aspartate hydrogen bond dyads (His142/Asp176 and His143/Asp183) are located near the bound acetyl lysine moiety (Figure 1.6). The enzyme is proposed to catalyze hydrolysis using a metal-coordinated water nucleophile and general acid-base catalysis (GABC) with either one or two side chains, similar to typical metallohydrolase mechanisms[16, 21, 103, 104] (Figure 1.6). The substrate binds to HDAC8 with the catalytic metal coordinating both the carbonyl oxygen of the acetyl lysine substrate and a water molecule. In the first step of the mechanism, His142 functions as a general base to abstract a proton from the metal-bound water that is reacting with the carbonyl carbon to form a high energy tetrahedral intermediate. The oxyanion intermediate is proposed to be stabilized by coordination with the metal ion, hydrogen bonding with Tyr306, and electrostatic interactions with positively charged groups in the active site. Proton donation from an active site general acid to the amine leaving group accompanies breakdown of the tetrahedral intermediate to form the acetate and the deacetylated lysine products[103]. In the GABC mechanism originally proposed from the crystal structure of the homologous HDLP enzyme[103], His142 and protonated His143 are proposed to function as the general base and general acid, respectively. In the one GABC mechanism, H143 functions as both the general acid and general base catalyst and H142 acts as an electrostatic catalyst[16, 104], similar to the

mechanism proposed for carboxypeptidase A[21]. Subsequent studies utilizing mutagenesis and molecular dynamics simulations suggest a preference for the one base mechanism[16, 44, 104].



**Figure 1.6 Schematic of the one base mechanism for HDAC8**

The acetyl lysine of the substrate is shown in blue while the nucleophilic water is green and red. For clarity, equilibration of exchangeable protons with solvent is not shown.

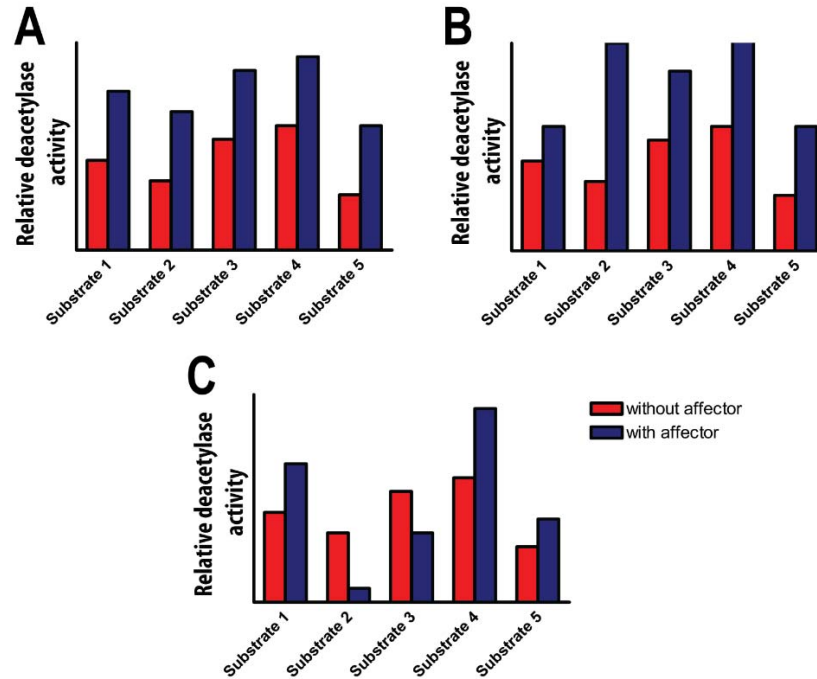
The HDAC8 crystal structure also contains two monovalent cation sites[13-18], suggesting that the activity of HDAC8 may be modulated by both the concentration and type of ions in solution. One monovalent cation site is 7 Å from the divalent catalytic metal ion and is coordinated by the side chain oxygens of Asp176 and Ser199 and the backbone carbonyl oxygens of Asp176, Asp178, His180, and Leu200. The second site is 21 Å from the divalent catalytic metal ion, and is ligated by two water molecules and the backbone carbonyl oxygens of Phe189, Thr192, Val195, and Tyr225. Initial activity measurements demonstrated that the concentrations of  $K^+$

and  $\text{Na}^+$  modulate HDAC8 catalysis[biolmol unpublished]. A detailed examination demonstrated that the value of  $k_{\text{cat}}/K_{\text{M}}$  for HDAC8-catalyzed deacetylation has a biphasic dependence on the concentration of  $\text{K}^+$  and  $\text{Na}^+$  ions[54]. In the absence of monovalent ions, the activity of HDAC8 is very low; addition of monovalent cations to Zn-bound HDAC8 increases activity with  $K_{1/2,\text{act}} = 14 \text{ mM}$  for  $\text{K}^+$ . At higher  $\text{K}^+$  concentrations Zn-HDAC8 activity is inhibited with  $K_{1/2,\text{inhib}} = 130 \text{ mM}$ . Mutagenesis studies indicate a significant decrease in potassium inhibition in the His142Ala and Asp176Ala/Asn mutants indicating that the monovalent ion site near the active site is inhibitory. Potassium binding next to His142 has been proposed to lower the  $\text{pK}_a$  of this residue, decreasing the concentration of protonated His142, thereby lowering catalytic activity. Similar biphasic regulation has been measured for  $\text{Na}^+$ , but activation and inhibition require a five-fold and ten-fold higher concentration of  $\text{Na}^+$  compared to  $\text{K}^+$ , respectively[54]. At the 100 mM  $\text{K}^+$  concentration within smooth muscle cells[105], HDAC8 activity is partially inhibited and sensitive to changes in the  $\text{K}^+$  concentration.

HDAC8 catalytic activity is enhanced by a number of divalent metal ions, including  $\text{Co}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Ni}^{2+}$ , and  $\text{Fe}^{2+}$ [26]. When HDAC8 is purified under aerobic conditions, the bound metal ion is  $\text{Zn}^{2+}$ . However, recombinant HDAC8 purified anaerobically from *E. coli* contains 8-fold more iron than zinc, consistent with this, the recombinant HDAC8 activity in *E. coli* cell lysates is oxygen-sensitive[26]. Additionally, although HDAC8 binds  $\text{Zn}^{2+}$  nearly  $10^6$ -fold more tightly than  $\text{Fe}^{2+}$ [17], the affinities for both metal ions are comparable to the readily exchangeable metal concentrations estimated in living cells, suggesting that HDAC8 can bind either  $\text{Fe}^{2+}$  or  $\text{Zn}^{2+}$  *in vivo*. Furthermore, the identity of the bound metal ion alters the catalytic properties of HDAC8. When catalyzing deacetylation of the methylcoumarin-labeled p53 peptide, the  $k_{\text{cat}}/K_{\text{M}}$  value for  $\text{Fe}^{2+}$ -bound HDAC8 is almost three times larger than that of  $\text{Zn}^{2+}$ -HDAC8. Interestingly, substitution of  $\text{Fe}^{2+}$  for  $\text{Zn}^{2+}$  also decreases the  $K_{\text{I}}$  for SAHA, suggesting that  $\text{Fe}^{2+}$  enhances ligand affinity[26]. However, a comparison of the crystal structures of the hydroxamate-bound  $\text{Fe}^{2+}$ -HDAC8 and  $\text{Zn}^{2+}$ -HDAC8 shows no significant differences in the active site or the rest of the protein[17]. These data suggest that either binding of the hydroxamic inhibitor stabilizes a common enzyme conformation, or that the bound metal ion affects protein dynamics that are not observable by crystallography.

Comparison of the  $Zn^{2+}/Fe^{2+}$  metal affinities with the cellular concentrations of those metals suggests that HDAC8 likely binds a combination of iron and zinc cofactors in eukaryotic cells[17]. Furthermore, the cellular zinc concentration can change dramatically upon oxidative stress[106, 107] and metal toxicity[108] potentially altering the populations of  $Fe^{2+}$ -HDAC8 and the  $Zn^{2+}$ -HDAC8 based on cellular conditions. This provides a means by which the cell could couple HDAC8 activity to cellular stresses.

A simple model for HDAC activation and inhibition assumes that compounds, cofactors, and binding partners equally affect the activity of HDAC8 with all substrates (Figure 1.7A). An alternative to this model proposes that substrate selectivity may be differentially regulated by stimuli. For example, scaffolding activators could preferentially enhance the binding of HDAC8 to one set of substrates (Figure 1.7B). Similarly, alteration of the active site metal ion or bound monovalent ions could alter ligand specificity. For example,  $Fe^{2+}$ -HDAC8 binds the inhibitor SAHA 2-fold more tightly than  $Zn^{2+}$ -HDAC8[54] even though  $Zn^{2+}$  is a stronger Lewis acid[109]. This change in binding affinity suggests that the active site metal ion may contribute subtly to the structure, dynamics, and molecular recognition of HDACs.



**Figure 1.7 Schematic of three potential models for describing the effect of an activating effector on HDAC activity**

A) In this model catalysis of deacetylation of each substrate is enhanced by an equivalent factor upon addition of the effector. B) In this model catalysis of deacetylation of each substrate is enhanced by a different factor upon addition of the effector. C) In this model catalysis of deacetylation of some substrates is activated while other substrates are inhibited by the effector.

### HDAC8 localization

Most simply, protein localization may regulate HDAC8 substrate specificity by changing the effective substrate concentration. HDACs have been found to have a range of cellular locations. HDAC1 and 2 are exclusively nuclear, while HDAC6 is mostly cytoplasmic, and HDAC3, 4, 5, 6, 7, 9, 10, and 11 appear to shuttle in and out of the nucleus[110]. Initially, HDAC8 was found to have a putative nuclear localization site and was observed in the nucleus of NIH3T3[27] and HEK293 cells[28]. Soon after, microscopy demonstrated that HDAC8 localizes to both the cytoplasm and nucleus of embryonic smooth muscle cells, skin fibroblasts, NIH3T3 cells, liver tissue[42, 50], and epithelial cells[79]. HDAC3, the closest HDAC8 human homologue[20], exists in both the cytoplasm and nucleus, and localization has been linked to the regulation and

cellular function of this enzyme. Whether cellular localization plays a role in HDAC8 activity is currently unknown, as no studies have yet broached this subject.

Determining the cell type-dependent expression of HDAC8 may provide interesting insights about its substrate specificity and biological function. In general, class I HDACs are ubiquitously expressed amongst the various cells of an organism, while class II HDACs are more cell-type specific[110, 111]. Likewise, HDAC8 has been found in a number of different healthy and diseased cell types.

HDAC8 knockouts after birth are non-lethal[112], consistent with the ability of humans to tolerate pan-HDAC inhibitors as an anti-cancer treatment[113]. However, protein expression profiles can vary significantly during development and several HDAC knockouts are lethal during mammalian embryonic development[8]. For example, cells lacking HDAC3 die before embryonic day 9.5; deletion of HDAC3 leads to hyperactivity of the nuclear receptor PPAR $\alpha$  and problems with embryonic gastrulation[114]. Similarly, HDAC8 expression is crucial to development, as mice lacking this enzyme die soon after birth[112]. Death is due to brain hemorrhaging caused by defects in the development of the mouse skull resulting from problems with neural crest patterning. Skull defects in mice are similar to those that occur upon overexpression of the transcription factors Otx2 and Lhx1, suggesting that HDAC8 either directly regulates these proteins or affects regulators of these proteins[112]. The mechanism of HDAC8 regulation of Otx2 and Lhx1 has yet to be determined, though treatment with an HDAC8 inhibitor suggests that HDAC8 may alter histone H3 acetylation levels in the promoter region of Otx2 and Lhx1[39]. Furthermore, since HDAC8 knockouts in mice are not lethal after birth[112], it is unclear whether HDAC8 is not essential for regulation of proteins, whether this regulation still occurs but is not vital for viability, or whether another mechanism is utilized. In humans some HDAC8 mutations and truncations are non-lethal but cause facial abnormalities and intellectual disabilities[68, 69, 115]. Because some HDAC8 mutants are not lethal in humans while still able to cause facial abnormalities could in part reflect to HDAC8's ability to act as a scaffold[61, 83]. Alternatively, HDAC8 could act by different mechanisms in mice and humans causing decreased lethality of these mutations.

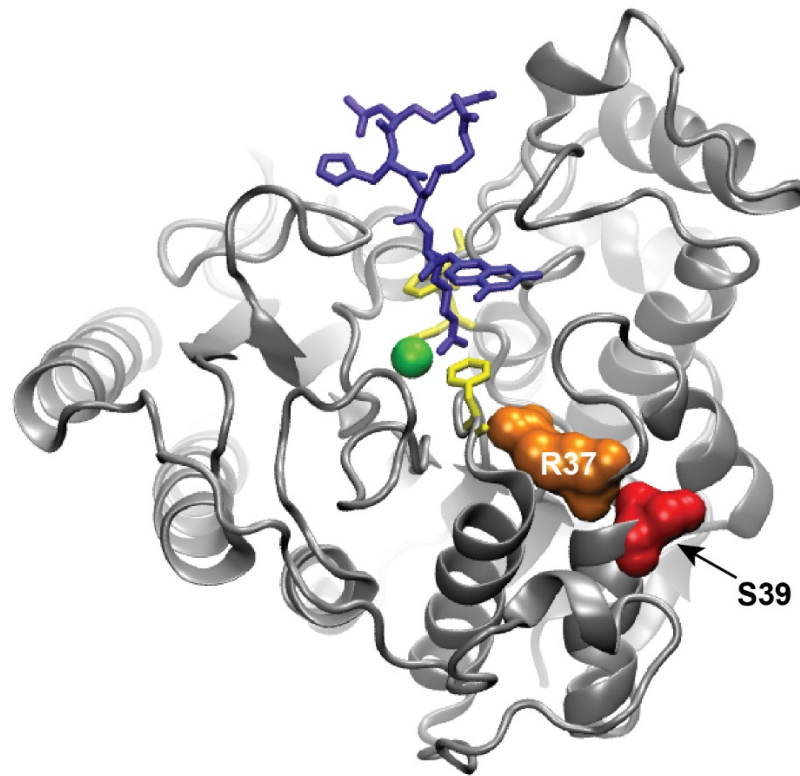
## Post-translational modification of HDAC8

Post-translational modifications such as phosphorylation, may also regulate HDAC8 activity. A screen of three protein kinases, casein kinase II, protein kinase A (PKA) and protein kinase G (PKG) indicated that phosphorylation of HDAC8 could be catalyzed by both PKA and PKG[116]. Phosphorylation of HDAC8 by PKA appeared to be predominant means by which HDAC8 is phosphorylated. This function was authenticated *in vivo* by incubation of cells with the PKA inhibitor H-89, which lowered HDAC8 phosphorylation levels[117]. Based on consensus sequences, nineteen potential phosphorylation sites were identified in HDAC8. Phosphoamino acid analysis followed by two dimensional thin layer chromatography demonstrated modification of a serine residue[117] and, based on this information, Ser39 was identified as the only PKA phosphorylation site in the HDAC8 sequence[27, 117]. A Ser39Ala HDAC8 mutant, which cannot be phosphorylated, negates phosphorylation of HDAC8 catalyzed by PKA, confirming this location as the primary phosphorylation site on HDAC8. Furthermore, phosphorylation of this site modulates HDAC8 activity. The specific activity of HDAC8 purified from cells treated with forskolin, a PKA activator, decreased by five-fold in an *in vitro* assay using purified histones[117]. Furthermore, the specific activity of Ser39Glu HDAC8, a mutation that mimics phosphorylation, decreases to a level comparable to that of phosphorylated HDAC8, while the specific activity of the Ser39Ala mutant is similar to unmodified HDAC8. To examine whether *in vivo* effects of phosphorylation of HDAC8 correlate with the *in vitro* measurements, HDAC8-transfected HeLa cells were treated with forskolin. These cells showed increased levels of acetylated histones H3 and H4, suggesting that the decreased deacetylase activity of phosphorylated HDAC8 led to increased acetylation *in vivo*[117].

Ser39 is located on the backside of the HDAC8 surface, 21Å from the catalytic metal ion[13-18] (Figure 1.8). Nonetheless, phosphorylation has the potential to affect the subcellular localization, protein-protein interactions, allosteric effects, and HDAC8 activity via conformational changes that propagate to the active site or enzyme-substrate interface. Ser39 lies near the junction with the L1 loop[13-18] that has been implicated in substrate recognition, and therefore phosphorylation at that position may alter enzyme-substrate interactions. The Ser39 residue is



located in a pocket on the enzyme surface surrounded by hydrophobic and acidic residues suggesting that phosphorylation of Ser39 could induce a drastic structural perturbation due to charge repulsion[14], where a Ser39Glu mutation decreases HDAC8 catalyzed deacetylation of core histones. Activity towards Ser39 also directly contacts the conserved Arg37 residue, which is proposed to be important for gating an acetate release channel in HDAC8[70] (Figure 1.8). The Arg37Ala mutation decreases the  $k_{cat}/K_M$  value for  $\text{Co}^{2+}$ -HDAC8-catalyzed deacetylation of the Fluor-de-lys substrate (R-H-Kac-Kac-fluorophore) by 530-fold[70]. Based on the proximity of Ser39 to Arg37, phosphorylation at this position may similarly affect HDAC8 activity.



**Figure 1.8 Phosphorylation of Ser39 may affect the active site structure and/or reactivity of HDAC8**

PDBID: 2v5w[13]. This structure shows that phosphorylation of Ser39 (red) may be able to perturb the position and/or electrostatic environment of Arg37 (orange) and in turn, affect the active site residues (yellow). Blue is the Fluor de Lys substrate and green is the active site metal.

Phosphorylation may also regulate HDAC8 through the modulation of protein-protein interactions. In the bacterial two-hybrid assay that identified fifteen HDAC8-interacting proteins[83] expression of PKA was necessary for the pulldown of six of these identified proteins, and suggests that these proteins interact solely with phosphoHDAC8. Two of these interactions, those between HDAC8 and hEST1B and between HDAC8 and Hsp70, were further

observed by co-immunoprecipitation, showing that treatment of cells with forskolin, led to increased amounts of phosphorylated HDAC8 and increased interactions[83]. These data strongly suggest that HDAC8 phosphorylation regulates HDAC8 complex formation as well as catalytic activity. Similarly, phosphorylation of HDAC1 and HDAC2 regulates association of these proteins with complexes such as mSin3A, RbAp48, and CoREST[116, 118].

Phosphorylation-dependent complex formation may also regulate the cellular localization of HDAC8. Fluorescence microscopy of myometrial cells shows that HDAC8 and phosphoHDAC8 both localize primarily to the cytosol, but cell fractionation data suggest that phosphoHDAC8 has increased association with the cytoskeleton compared to HDAC8 in this cell type[57].

HDAC4, HDAC5, and HDAC7 have been proposed to utilize nuclear-cytoplasmic shuttling mechanisms involving phosphorylation-dependent binding to 14-3-3 proteins for regulating their subcellular localization, and a similar mechanism may regulate HDAC8 localization[119-122].

The Ser39 site is an interesting location for phosphorylation amongst HDACs. Ser39 is not conserved among class I HDACs; the residue in the corresponding position of other class I HDACs is arginine in HDAC1 and 2, and alanine in HDAC3. Also, HDAC8 and HDAC5 contain the only phosphorylation sites that are located within the HDAC catalytic domain[123-125]. Additionally, HDAC8 is the only isozyme phosphorylated by PKA[125]. In general, the effect of phosphorylation on the activity of other class I isozymes, HDAC1 and 2, is ambiguous and/or contradictory[116, 118, 126, 127]. For example, phosphorylation of HDAC1 had little to no effect on deacetylase activity using a synthetic histone H4 peptide[126, 127] but activity on isolated histones decreased using mutants that could not be phosphorylated[118]. Therefore HDAC8 may be the best isozyme for examining the role of phosphorylation in regulating acetylation.

Many HDACs undergo additional post-translational modification including acetylation, ubiquitination, and sumoylation[124], but additional modifications of HDAC8 have not yet been demonstrated. HDAC8 has a consensus motif for glycosylation at Asn136 that could be modified[27, 117]; however the NetNGlyc 1.0 server does not predict N-glycosylation of this site due to the lack of a signal peptide[128]. Acetylation has been observed for HDAC1 at multiple sites, and one of the acetylated residues is conserved in HDAC8. Two of the HDAC1

sites are located in the deacetylase domain and four sites are near the C-terminus; acetylation of these sites inhibits HDAC1 deacetylase activity toward histones *in vitro* and corepressor function *in vivo*[129]. The two sites in the deacetylase domain, Lys218 and Lys220, are located near the activating monovalent cation binding site, so decreased activity from acetylation of these residues may arise from alteration of monovalent cation binding[129]. Sequence alignment by Cobalt indicates that the Lys218 position in HDAC1 is conserved in the corresponding Lys221 position in HDAC8(<http://www.ncbi.nlm.nih.gov/tools/cobalt/>). This lysine is within the proximity of monovalent site. As this site hypothetically activates HDAC8 allosterically[54], it is feasible that HDAC8 activity could be regulated by modification at this location. However, no modifications at this site have yet been observed and post-translational modifications of HDAC8 need to be further examined.

### **Concluding remarks**

Due to the abundance and vital function of acetylation within the cell, enzymes that catalyze acetylation and deacetylation are regulated in a multitude of ways and on a number of time scales. One mechanism of regulating HDAC activity is changing the substrate preferences for these enzymes, which in turn affects cellular processes. These regulatory mechanisms may allow the cell to finely tune the substrate preference for many HDACs simultaneously by allowing the same stimuli to differentially alter the activity of each HDAC isozyme. Understanding the interplay between various stimuli and HDAC regulation will give us tremendous insight into the inner workings of cellular processes and the mechanisms of disease formation. Even though HDAC8 has been extensively studied, it is humbling to know the vast amounts of information that have yet to be determined regarding the cohort of HDAC8 substrates and binding partners, localization in the cell, and regulatory mechanisms. Therefore, even for the best-characterized HDAC, there are likely many factors that affect substrate recognition that have not yet been discovered. The dissection of these factors in the future will be tremendously important for understanding not only the cellular function of HDACs, but also cellular regulation by post-translational modifications.

## Acknowledgements

We thank Dr. Patrick O'Brien for his useful insight into the paper's contents, Tony Mustoe for his help with visualization of molecules, Mark Taylor, Felicia Gray, Alison Tebo, and members of the Fierke Lab for their critical reading of this paper and discussion. This material was supported, in whole or in part, by the National Institutes of Health Grants, NIGMS GM40602 (CAF), T32-GM-008353, and 5-T32-GM-008597. This material is based upon work supported by the National Science Foundation under Grant No. DGE0718128.

## Bibliography

- [1] X.J. Yang, E. Seto, The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men, *Nat Rev Mol Cell Biol*, 9 (2008) 206-218.
- [2] M.A. Glozak, N. Sengupta, X. Zhang, E. Seto, Acetylation and deacetylation of non-histone proteins, *Gene*, 363 (2005) 15-23.
- [3] J.L. Workman, R.E. Kingston, Alteration of nucleosome structure as a mechanism of transcriptional regulation, *Annu Rev Biochem*, 67 (1998) 545-579.
- [4] M.A. Glozak, E. Seto, Histone deacetylases and cancer, *Oncogene*, 26 (2007) 5420-5432.
- [5] C. Choudhary, C. Kumar, F. Gnad, M.L. Nielsen, M. Rehman, T.C. Walther, J.V. Olsen, M. Mann, Lysine acetylation targets protein complexes and co-regulates major cellular functions, *Science*, 325 (2009) 834-840.
- [6] S. Zhao, W. Xu, W. Jiang, W. Yu, Y. Lin, T. Zhang, J. Yao, L. Zhou, Y. Zeng, H. Li, Y. Li, J. Shi, W. An, S.M. Hancock, F. He, L. Qin, J. Chin, P. Yang, X. Chen, Q. Lei, Y. Xiong, K.L. Guan, Regulation of cellular metabolism by protein lysine acetylation, *Science*, 327 (2010) 1000-1004.
- [7] G.A. Khoury, R.C. Baliban, C.A. Floudas, Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database, *Scientific reports*, 1 (2011).
- [8] M. Haberland, R.L. Montgomery, E.N. Olson, The many roles of histone deacetylases in development and physiology: implications for disease and therapy, *Nature reviews. Genetics*, 10 (2009) 32-42.

- [9] A.G. Kazantsev, L.M. Thompson, Therapeutic application of histone deacetylase inhibitors for central nervous system disorders, *Nat Rev Drug Discov*, 7 (2008) 854-868.
- [10] D.C. Drummond, C.O. Noble, D.B. Kirpotin, Z. Guo, G.K. Scott, C.C. Benz, Clinical development of histone deacetylase inhibitors as anticancer agents, *Annu Rev Pharmacol Toxicol*, 45 (2005) 495-528.
- [11] S. Spiegel, S. Miltien, S. Grant, Endogenous modulators and pharmacological inhibitors of histone deacetylases in cancer therapy, *Oncogene*, 31 (2012) 537-551.
- [12] L.N. Pan, J. Lu, B. Huang, HDAC inhibitors: a potential new category of anti-tumor agents, *Cell Mol Immunol*, 4 (2007) 337-343.
- [13] A. Vannini, C. Volpari, P. Gallinari, P. Jones, M. Mattu, A. Carfi, R. De Francesco, C. Steinkuhler, S. Di Marco, Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex, *EMBO Rep*, 8 (2007) 879-884.
- [14] J.R. Somoza, R.J. Skene, B.A. Katz, C. Mol, J.D. Ho, A.J. Jennings, C. Luong, A. Arvai, J.J. Buggy, E. Chi, J. Tang, B.C. Sang, E. Verner, R. Wynands, E.M. Leahy, D.R. Dougan, G. Snell, M. Navre, M.W. Knuth, R.V. Swanson, D.E. McRee, L.W. Tari, Structural snapshots of human HDAC8 provide insights into the class I histone deacetylases, *Structure*, 12 (2004) 1325-1334.
- [15] L. Whitehead, M.R. Dobler, B. Radetich, Y. Zhu, P.W. Atadja, T. Claiborne, J.E. Grob, A. McRiner, M.R. Pancost, A. Patnaik, W. Shao, M. Shultz, R. Tichkule, R.A. Tommasi, B. Vash, P. Wang, T. Stams, Human HDAC isoform selectivity achieved via exploitation of the acetate release channel with structurally unique small molecule inhibitors, *Bioorg Med Chem*, 19 (2011) 4626-4634.
- [16] D.P. Dowling, S.L. Gantt, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structural studies of human histone deacetylase 8 and its site-specific variants complexed with substrate and inhibitors, *Biochemistry*, 47 (2008) 13554-13563.
- [17] D.P. Dowling, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structures of metal-substituted human histone deacetylase 8 provide mechanistic inferences on biological function, *Biochemistry*, 49 (2010) 5048-5056.
- [18] K.E. Cole, D.P. Dowling, M.A. Boone, A.J. Phillips, D.W. Christianson, Structural basis of the antiproliferative activity of largazole, a depsipeptide inhibitor of the histone deacetylases, *J Am Chem Soc*, 133 (2011) 12474-12477.
- [19] A. Vannini, C. Volpari, G. Filocamo, E.C. Casavola, M. Brunetti, D. Renzoni, P. Chakravarty, C. Paolini, R. De Francesco, P. Gallinari, C. Steinkuhler, S. Di Marco, Crystal

structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor, *Proc Natl Acad Sci U S A*, 101 (2004) 15064-15069.

[20] I.V. Gregorette, Y.M. Lee, H.V. Goodson, Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis, *J Mol Biol*, 338 (2004) 17-31.

[21] M. Hernick, C.A. Fierke, Zinc hydrolases: the mechanisms of zinc-dependent deacetylases, *Arch Biochem Biophys*, 433 (2005) 71-84.

[22] R.A. Frye, Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins, *Biochem Biophys Res Commun*, 273 (2000) 793-798.

[23] J.C. Milne, J.M. Denu, The Sirtuin family: therapeutic targets to treat diseases of aging, *Curr Opin Chem Biol*, 12 (2008) 11-17.

[24] K.L. Durst, B. Lutterbach, T. Kummalue, A.D. Friedman, S.W. Hiebert, The *inv(16)* fusion protein associates with corepressors via a smooth muscle myosin heavy-chain domain, *Mol Cell Biol*, 23 (2003) 607-619.

[25] B.J. Wilson, A.M. Tremblay, G. Deblois, G. Sylvain-Drolet, V. Giguere, An acetylation switch modulates the transcriptional activity of estrogen-related receptor alpha, *Mol Endocrinol*, 24 (2010) 1349-1358.

[26] S.L. Gantt, S.G. Gattis, C.A. Fierke, Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion, *Biochemistry*, 45 (2006) 6170-6178.

[27] E. Hu, Z. Chen, T. Fredrickson, Y. Zhu, R. Kirkpatrick, G.F. Zhang, K. Johanson, C.M. Sung, R. Liu, J. Winkler, Cloning and characterization of a novel human class I histone deacetylase that functions as a transcription repressor, *J Biol Chem*, 275 (2000) 15254-15264.

[28] J.J. Buggy, M.L. Sideris, P. Mak, D.D. Lorimer, B. McIntosh, J.M. Clark, Cloning and characterization of a novel human histone deacetylase, HDAC8, *Biochem J*, 350 Pt 1 (2000) 199-205.

[29] I. Van den Wyngaert, W. de Vries, A. Kremer, J. Neefs, P. Verhasselt, W.H. Luyten, S.U. Kass, Cloning and characterization of human histone deacetylase 8, *FEBS Lett*, 478 (2000) 77-83.

[30] B.C. Smith, J.M. Denu, Acetyl-lysine analog peptides as mechanistic probes of protein deacetylases, *J Biol Chem*, 282 (2007) 37256-37265.

- [31] D. Riester, C. Hildmann, S. Grunewald, T. Beckers, A. Schwienhorst, Factors affecting the substrate specificity of histone deacetylases, *Biochem Biophys Res Commun*, 357 (2007) 439-445.
- [32] Z.A. Gurard-Levin, K.A. Kilian, J. Kim, K. Bahr, M. Mrksich, Peptide arrays identify isoform-selective substrates for profiling endogenous lysine deacetylase activity, *ACS Chem Biol*, 5 (2010) 863-873.
- [33] Z.A. Gurard-Levin, M. Mrksich, The activity of HDAC8 depends on local and distal sequences of its peptide substrates, *Biochemistry*, 47 (2008) 6242-6250.
- [34] A. Dose, S. Liokatis, F.X. Theillet, P. Selenko, D. Schwarzer, NMR profiling of histone deacetylase and acetyl-transferase activities in real time, *ACS Chem Biol*, 6 (2011) 419-424.
- [35] H. Lee, N. Rezai-Zadeh, E. Seto, Negative regulation of histone deacetylase 8 activity by cyclic AMP-dependent protein kinase A, *Mol Cell Biol*, 24 (2004) 765-773.
- [36] S. Balasubramanian, J. Ramos, W. Luo, M. Sirisawad, E. Verner, J.J. Buggy, A novel histone deacetylase 8 (HDAC8)-specific inhibitor PCI-34051 induces apoptosis in T-cell lymphomas, *Leukemia*, 22 (2008) 1026-1034.
- [37] D.E. Olson, F.F. Wagner, T. Kaya, J.P. Gale, N. Aidoud, E.L. Davoine, F. Lazzaro, M. Weiwer, Y.L. Zhang, E.B. Holson, Discovery of the first histone deacetylase 6/8 dual inhibitors, *J Med Chem*, 56 (2013) 4816-4820.
- [38] K. Krennhrubec, B.L. Marshall, M. Hedglin, E. Verdin, S.M. Ulrich, Design and evaluation of 'Linkerless' hydroxamic acids as selective HDAC8 inhibitors, *Bioorg Med Chem Lett*, 17 (2007) 2874-2878.
- [39] A. Saha, G.N. Pandian, S. Sato, J. Taniguchi, K. Hashiya, T. Bando, H. Sugiyama, Synthesis and biological evaluation of a targeted DNA-binding transcriptional activator with HDAC8 inhibitory activity, *Bioorg Med Chem*, 21 (2013) 4201-4209.
- [40] L.J. Juan, W.J. Shia, M.H. Chen, W.M. Yang, E. Seto, Y.S. Lin, C.W. Wu, Histone deacetylases specifically down-regulate p53-dependent gene activation, *J Biol Chem*, 275 (2000) 20436-20443.
- [41] J. Luo, F. Su, D. Chen, A. Shiloh, W. Gu, Deacetylation of p53 modulates its effect on cell growth and apoptosis, *Nature*, 408 (2000) 377-381.
- [42] D. Waltregny, L. De Leval, W. Glenisson, S. Ly Tran, B.J. North, A. Bellahcene, U. Weidle, E. Verdin, V. Castronovo, 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*, 165 (2004) 553-564.

[43] D. Waltregny, W. Glenisson, S.L. Tran, B.J. North, E. Verdin, A. Colige, V. Castronovo, Histone deacetylase HDAC8 associates with smooth muscle alpha-actin and is essential for smooth muscle cell contractility, *FASEB J*, 19 (2005) 966-968.

[44] S.L. Gantt, *Human Histone Deacetylase 8: Metal Dependence and Catalytic Mechanism*, Chemistry, The University of Michigan, Ann Arbor, 2006.

[45] A. Fersht, *Structure and mechanism in protein science : a guide to enzyme catalysis and protein folding*, W.H. Freeman, New York, 1999.

[46] D. Wegener, F. Wirsching, D. Riester, A. Schwienhorst, A fluorogenic histone deacetylase assay well suited for high-throughput activity screening, *Chemistry & Biology*, 10 (2003) 61-68.

[47] Z.A. Gurard-Levin, J. Kim, M. Mrksich, Combining mass spectrometry and peptide arrays to profile the specificities of histone deacetylases, *Chembiochem : a European journal of chemical biology*, 10 (2009) 2159-2161.

[48] B.E. Schultz, S. Misialek, J. Wu, J. Tang, M.T. Conn, R. Tahilramani, L. Wong, Kinetics and comparative reactivity of human class I and class IIb histone deacetylases, *Biochemistry*, 43 (2004) 11083-11091.

[49] D. Riester, D. Wegener, C. Hildmann, A. Schwienhorst, Members of the histone deacetylase superfamily differ in substrate specificity towards small synthetic substrates, *Biochem Biophys Res Commun*, 324 (2004) 1116-1123.

[50] J. Wu, C. Du, Z. Lv, C. Ding, J. Cheng, H. Xie, L. Zhou, S. Zheng, The up-regulation of histone deacetylase 8 promotes proliferation and inhibits apoptosis in hepatocellular carcinoma, *Dig Dis Sci*, 58 (2013) 3545-3553.

[51] A. Bar-Even, E. Noor, Y. Savir, W. Liebermeister, D. Davidi, D.S. Tawfik, R. Milo, The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters, *Biochemistry*, 50 (2011) 4402-4410.

[52] R. Alarcon, M.S. Orellana, B. Neira, E. Uribe, J.R. Garcia, N. Carvajal, Mutational analysis of substrate recognition by human arginase type I--agmatinase activity of the N130D variant, *FEBS J*, 273 (2006) 5625-5631.

[53] J.A. Villena, A. Kralli, ERRalpha: a metabolic function for the oldest orphan, *Trends Endocrinol Metab*, 19 (2008) 269-276.

- [54] S.L. Gantt, C.G. Joseph, C.A. Fierke, Activation and inhibition of histone deacetylase 8 by monovalent cations, *J Biol Chem*, 285 (2010) 6036-6043.
- [55] F. Mitelman, S. Heim, Quantitative acute leukemia cytogenetics, *Genes Chromosomes Cancer*, 5 (1992) 57-66.
- [56] P. Liu, S.A. Tarle, A. Hajra, D.F. Claxton, P. Marlton, M. Freedman, M.J. Siciliano, F.S. Collins, Fusion between transcription factor CBF beta/PEBP2 beta and a myosin heavy chain in acute myeloid leukemia, *Science*, 261 (1993) 1041-1044.
- [57] M. Karolczak-Bayatti, M. Sweeney, J. Cheng, L. Edey, S.C. Robson, S.M. Ulrich, A. Treumann, M.J. Taggart, G.N. Europe-Finner, Acetylation of heat shock protein 20 (Hsp20) regulates human myometrial activity, *J Biol Chem*, 286 (2011) 34346-34355.
- [58] K.T. Smith, S.A. Martin-Brown, L. Florens, M.P. Washburn, J.L. Workman, Deacetylase inhibitors dissociate the histone-targeting ING2 subunit from the Sin3 complex, *Chem Biol*, 17 (2010) 65-74.
- [59] M.Y. Balkhi, A.K. Trivedi, M. Geletu, M. Christopeit, S.K. Bohlander, H.M. Behre, G. Behre, Proteomics of acute myeloid leukaemia: Cytogenetic risk groups differ specifically in their proteome, interactome and post-translational protein modifications, *Oncogene*, 25 (2006) 7041-7058.
- [60] Q. Lu, A.E. Hutchins, C.M. Doyle, J.R. Lundblad, R.P. Kwok, Acetylation of cAMP-responsive element-binding protein (CREB) by CREB-binding protein enhances CREB-dependent transcription, *J Biol Chem*, 278 (2003) 15727-15734.
- [61] J. Gao, B. Siddoway, Q. Huang, H. Xia, Inactivation of CREB mediated gene transcription by HDAC8 bound protein phosphatase, *Biochem Biophys Res Commun*, 379 (2009) 1-5.
- [62] L.F. Michael, H. Asahara, A.I. Shulman, W.L. Kraus, M. Montminy, The phosphorylation status of a cyclic AMP-responsive activator is modulated via a chromatin-dependent mechanism, *Mol Cell Biol*, 20 (2000) 1596-1603.
- [63] S. Henikoff, J.G. Henikoff, Amino acid substitution matrices from protein blocks, *Proc Natl Acad Sci U S A*, 89 (1992) 10915-10919.
- [64] M. Bollen, Combinatorial control of protein phosphatase-1, *Trends Biochem Sci*, 26 (2001) 426-431.
- [65] V. Leone, G. Mansueto, G.M. Pierantoni, M. Tornincasa, F. Merolla, A. Cerrato, M. Santoro, M. Grieco, A. Scaloni, A. Celetti, A. Fusco, CCDC6 represses CREB1 activity by recruiting histone deacetylase 1 and protein phosphatase 1, *Oncogene*, 29 (2010) 4341-4351.

- [66] G. Canettieri, I. Morantte, E. Guzman, H. Asahara, S. Herzig, S.D. Anderson, J.R. Yates, 3rd, M. Montminy, Attenuation of a phosphorylation-dependent activator by an HDAC-PP1 complex, *Nat Struct Biol*, 10 (2003) 175-181.
- [67] L. Mannini, F. Cucco, V. Quarantotti, I.D. Krantz, A. Musio, Mutation spectrum and genotype-phenotype correlation in Cornelia de Lange syndrome, *Hum Mutat*, 34 (2013) 1589-1596.
- [68] F.J. Kaiser, M. Ansari, D. Braunholz, M. Concepcion Gil-Rodriguez, C. Decroos, J.J. Wilde, C.T. Fincher, M. Kaur, M. Bando, D.J. Amor, P.S. Atwal, M. Bahlo, C.M. Bowman, J.J. Bradley, H.G. Brunner, D. Clark, M. Del Campo, N. Di Donato, P. Diakumis, H. Dubbs, D.A. Dymont, J. Eckhold, S. Ernst, J.C. Ferreira, L.J. Francey, U. Gehlken, E. Guillen-Navarro, Y. Gyftodimou, B.D. Hall, R. Hennekam, L. Hudgins, M. Hullings, J.M. Hunter, H. Yntema, A.M. Innes, A.D. Kline, Z. Krumina, H. Lee, K. Leppig, S.A. Lynch, M.B. Mallozzi, L. Mannini, S. McKee, S.G. Mehta, I. Micule, S. Mohammed, E. Moran, G.R. Mortier, J.A. Moser, S.E. Noon, N. Nozaki, L. Nunes, J.G. Pappas, L.S. Penney, A. Perez-Aytes, M.B. Petersen, B. Puisac, N. Revencu, E. Roeder, S. Saitta, A.E. Scheuerle, K.L. Schindeler, V.M. Siu, Z. Stark, S.P. Strom, H. Thiese, I. Vater, P. Willems, K. Williamson, L.C. Wilson, H. Hakonarson, F. Quintero-Rivera, J. Wierzba, A. Musio, G. Gillessen-Kaesbach, F.J. Ramos, L.G. Jackson, K. Shirahige, J. Pie, D.W. Christianson, I.D. Krantz, D.R. Fitzpatrick, M.A. Deardorff, Loss-of-function HDAC8 mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance, *Hum Mol Genet*, (2014).
- [69] M.A. Deardorff, M. Bando, R. Nakato, E. Watrin, T. Itoh, M. Minamino, K. Saitoh, M. Komata, Y. Katou, D. Clark, K.E. Cole, E. De Baere, C. Decroos, N. Di Donato, S. Ernst, L.J. Francey, Y. Gyftodimou, K. Hirashima, M. Hullings, Y. Ishikawa, C. Jaulin, M. Kaur, T. Kiyono, P.M. Lombardi, L. Magnaghi-Jaulin, G.R. Mortier, N. Nozaki, M.B. Petersen, H. Seimiya, V.M. Siu, Y. Suzuki, K. Takagaki, J.J. Wilde, P.J. Willems, C. Prigent, G. Gillessen-Kaesbach, D.W. Christianson, F.J. Kaiser, L.G. Jackson, T. Hirota, I.D. Krantz, K. Shirahige, HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle, *Nature*, 489 (2012) 313-317.
- [70] S. Haider, C.G. Joseph, S. Neidle, C.A. Fierke, M.J. Fuchter, On the function of the internal cavity of histone deacetylase protein 8: R37 is a crucial residue for catalysis, *Bioorg Med Chem Lett*, 21 (2011) 2129-2132.
- [71] N. Sengupta, E. Seto, Regulation of histone deacetylase activities, *J Cell Biochem*, 93 (2004) 57-67.
- [72] T. Hayakawa, J. Nakayama, Physiological roles of class I HDAC complex and histone demethylase, *J Biomed Biotechnol*, 2011 (2011) 129383.
- [73] Y. Qian, J. Zhang, Y.S. Jung, X. Chen, DEC1 coordinates with HDAC8 to differentially regulate TAp73 and DeltaNp73 expression, *PLoS One*, 9 (2014) e84015.

- [74] C.Q. Chen, K. Yu, Q.X. Yan, C.Y. Xing, Y. Chen, Z. Yan, Y.F. Shi, K.W. Zhao, S.M. Gao, Pure curcumin increases the expression of SOCS1 and SOCS3 in myeloproliferative neoplasms through suppressing class I histone deacetylases, *Carcinogenesis*, 34 (2013) 1442-1449.
- [75] S.M. Gao, C.Q. Chen, L.Y. Wang, L.L. Hong, J.B. Wu, P.H. Dong, F.J. Yu, Histone deacetylases inhibitor sodium butyrate inhibits JAK2/STAT signaling through upregulation of SOCS1 and SOCS3 mediated by HDAC8 inhibition in myeloproliferative neoplasms, *Exp Hematol*, 41 (2013) 261-270 e264.
- [76] E. Appella, C.W. Anderson, Post-translational modifications and activation of p53 by genotoxic stresses, *Eur J Biochem*, 268 (2001) 2764-2772.
- [77] C.M. Dreiza, C.M. Brophy, P. Komalavilas, E.J. Furnish, L. Joshi, M.A. Pallero, J.E. Murphy-Ullrich, M. von Rechenberg, Y.S. Ho, B. Richardson, N. Xu, Y. Zhen, J.M. Peltier, A. Panitch, Transducible heat shock protein 20 (HSP20) phosphopeptide alters cytoskeletal dynamics, *FASEB J*, 19 (2005) 261-263.
- [78] M. Vicente-Manzanares, X. Ma, R.S. Adelstein, A.R. Horwitz, Non-muscle myosin II takes centre stage in cell adhesion and migration, *Nat Rev Mol Cell Biol*, 10 (2009) 778-790.
- [79] Y. Yamauchi, H. Boukari, I. Banerjee, I.F. Sbalzarini, P. Horvath, A. Helenius, Histone deacetylase 8 is required for centrosome cohesion and influenza A virus entry, *PLoS Pathog*, 7 (2011) e1002316.
- [80] H.S. Chen, P. Wikramasinghe, L. Showe, P.M. Lieberman, Cohesins repress Kaposi's sarcoma-associated herpesvirus immediate early gene transcription during latency, *J Virol*, 86 (2012) 9454-9464.
- [81] T. Tanaka, J. Fuchs, J. Loidl, K. Nasmyth, Cohesin ensures bipolar attachment of microtubules to sister centromeres and resists their precocious separation, *Nat Cell Biol*, 2 (2000) 492-499.
- [82] R.W. Wong, G. Blobel, Cohesin subunit SMC1 associates with mitotic microtubules at the spindle pole, *Proc Natl Acad Sci U S A*, 105 (2008) 15441-15445.
- [83] H. Lee, N. Sengupta, A. Villagra, N. Rezai-Zadeh, E. Seto, Histone deacetylase 8 safeguards the human ever-shorter telomeres 1B (hEST1B) protein from ubiquitin-mediated degradation, *Mol Cell Biol*, 26 (2006) 5259-5269.
- [84] S. Koyasu, E. Nishida, T. Kadowaki, F. Matsuzaki, K. Iida, F. Harada, M. Kasuga, H. Sakai, I. Yahara, Two mammalian heat shock proteins, HSP90 and HSP100, are actin-binding proteins, *Proc Natl Acad Sci U S A*, 83 (1986) 8054-8058.

- [85] A. Taiyab, M. Rao Ch, HSP90 modulates actin dynamics: inhibition of HSP90 leads to decreased cell motility and impairs invasion, *Biochim Biophys Acta*, 1813 (2011) 213-221.
- [86] P. Joshi, T.M. Greco, A.J. Guise, Y. Luo, F. Yu, A.I. Nesvizhskii, I.M. Cristea, The functional interactome landscape of the human histone deacetylase family, *Molecular systems biology*, 9 (2013) 672.
- [87] J.J. Kovacs, C. Hubbert, T.P. Yao, The HDAC complex and cytoskeleton, *Novartis Foundation symposium*, 259 (2004) 170-177; discussion 178-181, 223-175.
- [88] I. Nobeli, A.D. Favia, J.M. Thornton, Protein promiscuity and its implications for biotechnology, *Nature biotechnology*, 27 (2009) 157-167.
- [89] L. Hedstrom, Serine protease mechanism and specificity, *Chem Rev*, 102 (2002) 4501-4524.
- [90] D.L. Ollis, E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S.M. Franken, M. Harel, S.J. Remington, I. Silman, J. Schrag, et al., The alpha/beta hydrolase fold, *Protein Eng*, 5 (1992) 197-211.
- [91] T.R. Weikl, C. von Deuster, Selected-fit versus induced-fit protein binding: kinetic differences and mutational analysis, *Proteins*, 75 (2009) 104-110.
- [92] G. Estiu, N. West, R. Mazitschek, E. Greenberg, J.E. Bradner, O. Wiest, On the inhibition of histone deacetylase 8, *Bioorg Med Chem*, 18 (2010) 4103-4110.
- [93] R.K. Singh, N. Lall, T.S. Leedahl, A. McGillivray, T. Mandal, M. Haldar, S. Mallik, G. Cook, D.K. Srivastava, Kinetic and thermodynamic rationale for suberoylanilide hydroxamic acid being a preferential human histone deacetylase 8 inhibitor as compared to the structurally similar ligand, trichostatin a, *Biochemistry*, 52 (2013) 8139-8149.
- [94] M. Brunsteiner, P.A. Petukhov, Insights from comprehensive multiple receptor docking to HDAC8, *J Mol Model*, 18 (2012) 3927-3939.
- [95] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, Origin of attraction and directionality of the pi/pi interaction: model chemistry calculations of benzene dimer interaction, *J Am Chem Soc*, 124 (2002) 104-112.
- [96] M.O. Sinnokrot, E.F. Valeev, C.D. Sherrill, Estimates of the ab initio limit for pi-pi interactions: the benzene dimer, *J Am Chem Soc*, 124 (2002) 10887-10893.

- [97] A. Horovitz, L. Serrano, B. Avron, M. Bycroft, A.R. Fersht, Strength and co-operativity of contributions of surface salt bridges to protein stability, *J Mol Biol*, 216 (1990) 1031-1044.
- [98] P.M. Lombardi, H.D. Angell, D.A. Whittington, E.F. Flynn, K.R. Rajashankar, D.W. Christianson, Structure of prokaryotic polyamine deacetylase reveals evolutionary functional relationships with eukaryotic histone deacetylases, *Biochemistry*, 50 (2011) 1808-1817.
- [99] J.C. Bressi, A.J. Jennings, R. Skene, Y. Wu, R. Melkus, R. De Jong, S. O'Connell, C.E. Grimshaw, M. Navre, A.R. Gangloff, Exploration of the HDAC2 foot pocket: Synthesis and SAR of substituted N-(2-aminophenyl)benzamides, *Bioorg Med Chem Lett*, 20 (2010) 3142-3145.
- [100] M.J. Bottomley, P. Lo Surdo, P. Di Giovine, A. Cirillo, R. Scarpelli, F. Ferrigno, P. Jones, P. Neddermann, R. De Francesco, C. Steinkuhler, P. Gallinari, A. Carfi, Structural and functional analysis of the human HDAC4 catalytic domain reveals a regulatory structural zinc-binding domain, *J Biol Chem*, 283 (2008) 26694-26704.
- [101] A. Schuetz, J. Min, A. Allali-Hassani, M. Schapira, M. Shuen, P. Loppnau, R. Mazitschek, N.P. Kwiatkowski, T.A. Lewis, R.L. Maglathin, T.H. McLean, A. Bochkarev, A.N. Plotnikov, M. Vedadi, C.H. Arrowsmith, Human HDAC7 harbors a class IIa histone deacetylase-specific zinc binding motif and cryptic deacetylase activity, *J Biol Chem*, 283 (2008) 11355-11363.
- [102] J.S. Papadopoulos, R. Agarwala, COBALT: constraint-based alignment tool for multiple protein sequences, *Bioinformatics*, 23 (2007) 1073-1079.
- [103] M.S. Finnin, J.R. Donigian, A. Cohen, V.M. Richon, R.A. Rifkind, P.A. Marks, R. Breslow, N.P. Pavletich, Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors, *Nature*, 401 (1999) 188-193.
- [104] R. Wu, S. Wang, N. Zhou, Z. Cao, Y. Zhang, A proton-shuttle reaction mechanism for histone deacetylase 8 and the catalytic role of metal ions, *J Am Chem Soc*, 132 (2010) 9471-9479.
- [105] S.E. Kasner, M.B. Ganz, Regulation of intracellular potassium in mesangial cells: a fluorescence analysis using the dye, PBFI, *Am J Physiol*, 262 (1992) F462-467.
- [106] C.J. Chang, J. Jaworski, E.M. Nolan, M. Sheng, S.J. Lippard, A tautomeric zinc sensor for ratiometric fluorescence imaging: application to nitric oxide-induced release of intracellular zinc, *Proc Natl Acad Sci U S A*, 101 (2004) 1129-1134.
- [107] R.A. Bozym, R.B. Thompson, A.K. Stoddard, C.A. Fierke, Measuring picomolar intracellular exchangeable zinc in PC-12 cells using a ratiometric fluorescence biosensor, *ACS Chem Biol*, 1 (2006) 103-111.

- [108] D. Wang, O. Hosteen, C.A. Fierke, ZntR-mediated transcription of zntA responds to nanomolar intracellular free zinc, *J Inorg Biochem*, (2012).
- [109] B.E. Douglas, D.H. McDaniel, J.J. Alexander, Concepts and models of inorganic chemistry, 3rd ed., Wiley, New York, 1994.
- [110] A.J. de Ruijter, A.H. van Gennip, H.N. Caron, S. Kemp, A.B. van Kuilenburg, Histone deacetylases (HDACs): characterization of the classical HDAC family, *Biochem J*, 370 (2003) 737-749.
- [111] X.J. Yang, S. Gregoire, Class II histone deacetylases: from sequence to function, regulation, and clinical implication, *Mol Cell Biol*, 25 (2005) 2873-2884.
- [112] M. Haberland, M.H. Mokalled, R.L. Montgomery, E.N. Olson, Epigenetic control of skull morphogenesis by histone deacetylase 8, *Genes Dev*, 23 (2009) 1625-1630.
- [113] K. Garber, HDAC inhibitors overcome first hurdle, *Nat Biotechnol*, 25 (2007) 17-19.
- [114] R.L. Montgomery, M.J. Potthoff, M. Haberland, X. Qi, S. Matsuzaki, K.M. Humphries, J.A. Richardson, R. Bassel-Duby, E.N. Olson, Maintenance of cardiac energy metabolism by histone deacetylase 3 in mice, *J Clin Invest*, 118 (2008) 3588-3597.
- [115] M. Harakalova, M.J. van den Boogaard, R. Sinke, S. van Lieshout, M.C. van Tuil, K. Duran, I. Renkens, P.A. Terhal, C. de Kovel, I.J. Nijman, M. van Haelst, N.V. Knoers, G. van Haften, W. Kloosterman, R.C. Hennekam, E. Cuppen, H.K. Ploos van Amstel, X-exome sequencing identifies a HDAC8 variant in a large pedigree with X-linked intellectual disability, truncal obesity, gynaecomastia, hypogonadism and unusual face, *J Med Genet*, 49 (2012) 539-543.
- [116] S.C. Tsai, Regulation of Histone deacetylase 2 by protein kinase CK2, *Journal of Biological Chemistry*, 277 (2002) 31826-31833.
- [117] H. Lee, N. Rezai-Zadeh, E. Seto, Negative regulation of histone deacetylase 8 activity by cyclic AMP-dependent protein kinase A, *Molecular and Cellular Biology*, 24 (2004) 765-773.
- [118] M.K. Pflum, J.K. Tong, W.S. Lane, S.L. Schreiber, Histone deacetylase 1 phosphorylation promotes enzymatic activity and complex formation, *The Journal of biological chemistry*, 276 (2001) 47733-47741.
- [119] A.H. Wang, M.J. Kruhlak, J. Wu, N.R. Bertos, M. Vezmar, B.I. Posner, D.P. Bazett-Jones, X.J. Yang, Regulation of histone deacetylase 4 by binding of 14-3-3 proteins, *Mol Cell Biol*, 20 (2000) 6904-6912.

- [120] C.M. Grozinger, 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 of the United States of America*, 97 (2000) 7835-7840.
- [121] X. Li, S. Song, Y. Liu, S.H. Ko, H.Y. Kao, Phosphorylation of the histone deacetylase 7 modulates its stability and association with 14-3-3 proteins, *The Journal of biological chemistry*, 279 (2004) 34201-34208.
- [122] H.Y. Kao, A. Verdel, C.C. Tsai, C. Simon, H. Juguilon, S. Khochbin, Mechanism for nucleocytoplasmic shuttling of histone deacetylase 7, *The Journal of biological chemistry*, 276 (2001) 47496-47507.
- [123] T.M. Greco, F. Yu, A.J. Guise, I.M. Cristea, Nuclear import of histone deacetylase 5 by requisite nuclear localization signal phosphorylation, *Molecular & Cellular Proteomics : MCP*, 10 (2011) M110 004317.
- [124] A. Brandl, T. Heinzel, O.H. Kramer, Histone deacetylases: salesmen and customers in the post-translational modification market, *Biology of the cell / under the auspices of the European Cell Biology Organization*, 101 (2009) 193-205.
- [125] E. Seto, X.-J. Yang, Regulation of Histone Deacetylase Activities and Functions by Phosphorylation and Dephosphorylation, in: R. Bradshaw, E. Dennis (Eds.) *Handbook of Cell Signaling*, Elsevier, Inc.2010, pp. 2379-2388.
- [126] S.C. Galasinski, K.A. Resing, J.A. Goodrich, N.G. Ahn, Phosphatase inhibition leads to histone deacetylases 1 and 2 phosphorylation and disruption of corepressor interactions, *The Journal of biological chemistry*, 277 (2002) 19618-19626.
- [127] R. Cai, P. Kwon, Y. Yan-Neale, L. Sambuccetti, D. Fischer, D. Cohen, Mammalian histone deacetylase 1 protein is posttranslationally modified by phosphorylation, *Biochemical and biophysical research communications*, 283 (2001) 445-453.
- [128] R. Gupta, E. Jung, S. Brunak, Prediction of N-glycosylation sites in human proteins, In preparation, (2004).
- [129] Z.Y. Qiu Y, Becker M, John S, Parekh BS, Huang S, Hendarwanto A, Martinez ED, Chen Y, Lu H, Adkins NL, Stavreva DA, Wiench M, Georgel PT, Schiltz RL, Hager GL., HDAC1 acetylation is linked to progressive modulation of steroid receptor-induced gene transcription, *Mol Cell*, 22 (2006) 669-679.



## Chapter 2

### **An enzyme-coupled assay measuring acetate production for profiling histone deacetylase specificity<sup>1,2</sup>**

#### **Introduction**

Histone (or acetyl lysine) deacetylases (HDACs) are a family of 18 enzymes that catalyze the deacetylation of acetylated lysine side chains[1, 2]. Acetylation is a post-translational modification identified on over 3,100 lysines within the mammalian proteome[3] that alter the activities and properties of modified proteins[4]. As many of these proteins are essential to cellular processes[5, 6], aberrant acetylation and deacetylation may contribute to disease states[7]. Attesting to the role of HDACs in diseases are two HDAC inhibitors (Vorinostat and Romidepsin) that have been approved by the FDA for the treatment of T-cell Lymphoma[8], though the mechanism of action for these drugs is not well understood. One complicating factor to understanding the biological function and regulation of protein deacetylation is the lack of identified HDAC isozyme-substrate pairs. Determining the substrate specificity of HDACs would provide insight into cellular homeostasis and development of isozyme-specific inhibitors.

---

<sup>1</sup> Reproduced in part from Wolfson, N.A. Pitcairn, C.A. Sullivan, E.D. Joseph, C.G. Fierke, C.A. 2014. Anal Biochem. 2014 Mar 24. pii: S0003-2697(14)00102-X. doi: 10.1016/j.ab.2014.03.012.

<sup>2</sup> Noah Wolfson developed the coupled acetate assay. Carol Ann Pitcairn developed the fluorescamine assay, acetylated tetramer and performed the assay on the tetramer. Eric Sullivan assayed HDAC3 and performed the continuous assay. Caleb Joseph developed the quench methodology.

There are four classes of HDAC enzymes. Classes I, II, and IV use an active site divalent metal ion cofactor to catalyze deacetylation, yielding lysine and acetate as products[9]. Class III HDACs use NAD<sup>+</sup> as a cofactor and produce 2'-O'acetylribose, nicotinamide, and the deacetylated protein[10]. Current *in vitro* assays for the measurement of HDAC activity use environmentally sensitive fluorophores[9, 11-17], HPLC methods[18], free amine reactive reagents[11], radiolabeled acetate[19, 20], and mass spectrometry[21-23]. While these assays are useful, each has associated limitations. In particular, many of these techniques can only be used to measure deacetylation of short peptide substrates rather than the biologically-relevant acetylated proteins. Furthermore, in the most frequently used assay, the Fluor de Lys assay, the methylcoumarin substituent alters substrate recognition[23]. For assays that can be adapted to measure deacetylation of proteins, many cannot be adapted to high throughput formats, are hard to quantify, or require specialized equipment.

Here we optimize an enzyme-coupled assay for the measurement of low micromolar acetate concentrations that can be used to evaluate the activity of class I, II, and IV HDACs. This assay couples the formation of acetate to the production of NADH that is monitored via an absorbance or fluorescence signal. This assay quantitatively measures deacetylation independent of the substrate size or structure, does not require specialized equipment, and can be adapted to real-time and high throughput formats. Using this assay, we demonstrate that cobalt(II)-bound histone deacetylase 8 (Co(II)-HDAC8) catalyzes deacetylation of a peptide sequence from the H4 histone tail containing a C-terminal methylcoumarin fluorophore with a  $k_{cat}/K_M$  value that is 50-fold, 2.8-fold, and 2.3-fold greater than the  $k_{cat}/K_M$  value for deacetylation of the same peptide

containing a C-terminal carboxylate, amide, or tryptophan, respectively. The loss of catalytic efficiency for deacetylation of the C-terminal carboxylate peptide is a result of a 2.3-fold increase in the value of  $K_M$  and a 22-fold decrease in the  $k_{cat}$  value. These data demonstrate that interactions between HDAC8 and the C-terminal moiety are important for substrate recognition and efficient chemistry.

## **Materials and Methods**

### **Reagents**

ATP, Coenzyme A,  $NAD^+$ , L-malic acid, citrate synthase (CS), and malate dehydrogenase (MDH) were purchased from Sigma. The acetic acid detection kit was purchased from R-biopharm. Fluor de Lys peptide and the developing reagent were purchased from Enzo Life Sciences. The unlabeled peptides (Ac-KGGAKac-COO<sup>-</sup>, Ac-KGGAKac-NH<sub>2</sub>, and Ac-KGGAKacW-NH<sub>2</sub>) were purchased from Peptide2.0 (>85 % purity). Cobalt and magnesium were purchased as ICP standards from GFS Chemicals and the acetic acid standard was purchased from the Ricca Chemical Company. Chelex 100 resin was purchased from Bio-Rad. HDAC3/NCOR1 was purchased from Enzo Life Sciences. Ethylenediaminetetraacetic acid (EDTA) was purchased from Sigma-Aldrich at >99 % purity. All other materials were purchased from Fisher and were of a purity >95 % unless otherwise noted.

### **Acetyl-CoA synthetase preparation**

The chitin tagged acetyl-CoA synthetase (ACS) plasmid (Acs/pTYB1)[24] was a generous gift from Professor Andrew Gulick (Hauptman-Woodward Institute). To increase the yield of

protein, the gene for ACS was subcloned from the chitin-tagged ACS plasmid into a pET vector containing a His<sub>6x</sub> affinity tag. The ACS gene from Acs/pTYB1 was amplified using the polymerase chain reaction to add XhoI and XbaI restriction sites. The amplified DNA segment was digested using XhoI and XbaI and ligated into a pHD4 vector[25] containing a T7 RNA polymerase promoter and a C-terminal TEV (Tobacco Etch Virus) protease sequence followed by a His<sub>6x</sub> motif to form the pHD4-ACS-TEV-His<sub>6x</sub> expression vector. The plasmid sequence was confirmed by sequencing at the University of Michigan DNA Sequencing Core.

The pHD4-ACS-TEV-His<sub>6x</sub> vector was transformed into BL21(DE3) cells. An overnight culture (12.5 mL/L) was used to inoculate autoinduction TB medium (12 g/L tryptone, 24 g/L yeast extract, 4.6 g/L KH<sub>2</sub>PO<sub>4</sub>, 20.6 g/L K<sub>2</sub>HPO<sub>4</sub>, 4 g/L lactose, 1 g/L glucose, 10 mL/L glycerol) that was supplemented with 100 µg/mL ampicillin and grown at 30°C for 20 hours prior to harvest. The cells were pelleted by centrifugation (9,000 x g, 10 min) and then resuspended and lysed in low imidazole buffer (30 mM HEPES, 150 mM NaCl, 20 mM imidazole, 1 mM TCEP, pH 8) using an M110L microfluidizer (Microfluidics). The lysate was cleared by centrifugation (39,000 x g, 45 min) and the supernatant was loaded onto a 12 mL GE Chelating Sepharose column charged with NiCl<sub>2</sub>. The ACS was eluted with a gradient of low (20 mM) to high (200 mM) imidazole buffer. Fractions containing ACS were identified using SDS-PAGE chromatography and were concentrated to <1 mL using 30,000 MWCO Amicon Ultra-15 centrifugal units and loaded onto a GE HiPrep 16/60 Sephacryl S200 HR size exclusion column equilibrated with size exclusion buffer (30 mM HEPES, 150 mM NaCl, 1 mM TCEP, pH 8). The fractions containing ACS were collected, combined with his-tagged TEV(S219V) protease (0.5 mg per liter of culture) purified in our lab using the method of Tropea *et al.*[26] and dialyzed against >500 fold

excess of low imidazole buffer at 4°C overnight. The dialyzed ACS was run over a second Ni<sup>2+</sup>-charged Sepharose column, and the flow-through containing ACS was collected and concentrated to ~2 mM. The protein was then flash frozen in liquid nitrogen and stored at -80°C. Frozen ACS can be used for at least 12 months with little to no effect on the assay.

### **HDAC8 expression and purification**

HDAC8 was expressed and purified as previously described[9] with the exception that a 20 mL DEAE Sepharose column was used after the second Chelating Sepharose column to remove excess metal from HDAC8. This column utilized a gradient from low to high salt buffer (50 mM HEPES, 10 μM ZnSO<sub>4</sub>, 1 mM TCEP, 50 mM NaCl, 5 mM KCl, pH 7.8 and buffer with 50 mM HEPES with 1 M NaCl).

### **Fluor de Lys assay**

All assays were performed in metal free tubes using metal free tips before being quenched into 96 well black plates (Corning plate# 3638). The Fluor de Lys assay was performed as previously described[9]. Briefly, HDAC8 was reconstituted with stoichiometric cobalt(II) at a final concentration of 10 μM and incubated on ice for 1 hour. Fluor de Lys HDAC8 deacetylase substrate (Ac-KGGAKac-methylcoumarin) (0 to 100 μM) was resuspended in HDAC8 assay buffer (50 mM HEPES, 137 mM NaCl, 2.7 mM KCl, pH 8) and incubated at 30°C for 5 minutes. Reactions were initiated by adding 0.5 μM Co(II)-HDAC8 and the reaction was quenched by a 10-fold dilution into 0.05x Enzo developer II and 1.2 μM Trichostatin A (TSA) in HDAC8 assay

buffer (above) at 0, 30, and 60 seconds. Samples were incubated at room temperature for 15 minutes and the fluorescence was measured using a Polarstar Galaxy fluorometer (ex. = 340 nm; em. = 450 nm and 380 nm). The initial rate of deacetylation was determined from the time-dependent increase in the fluorescence ratio (450 nm/380 nm) and the concentration of product was calculated using a standard curve.

### **Acetate assay kit**

Acetate standard curves were made by diluting the Ricca acetic acid standard with HDAC8 assay buffer (above). The acetic acid detection kit (R-biopharm) was used according to the instructions except that the reaction volume and coupled solution volume were decreased 10-fold and no additional water was added to dilute the reaction. Solutions 1, 2, 3 and 4 in the kit were preincubated at room temperature for 20 minutes before being mixed with acetate. The reaction was incubated at room temperature for 40 minutes and then NADH fluorescence (ex. = 340 nm, em. = 460 nm) was measured using a Polarstar Galaxy fluorometer in a 96 well plate.

### **Optimized stopped coupled acetate assay**

To remove contaminating metals from peptide substrates, ~6 % (v/v) hydrated Chelex 100 was added to the Ac-KGGAKac-NH<sub>2</sub> and Ac-KGGAKacW-NH<sub>2</sub> peptides and incubated at room temperature for three hours. The Ac-KGGAKacW-NH<sub>2</sub> peptide concentration was determined from the absorbance measurement (OD<sub>280</sub>) using an ND-1000 Spectrophotometer (Nanodrop) with a calculated extinction coefficient of 5500 M<sup>-1</sup> cm<sup>-1</sup>[27]. Additionally, the concentration of peptides containing a free amine (lysine) was measured using the fluorescamine assay described

below. Peptide substrates without a fluorophore (0 – 1600  $\mu\text{M}$ ) were preincubated in HDAC8 assay buffer (above) at 30°C for 10 min. The reactions were initiated by adding 0.5  $\mu\text{M}$  (final concentration) Co(II)-HDAC8 or HDAC3/NCOR1, and quenched by addition of 0.37 % (v/v, final concentration) HCl after 0, 30, 60, and 90 minutes of incubation. The reactions were flash frozen within 20 minutes of quenching and stored at -80°C. Upon thawing, the reactions were neutralized by addition of 0.6% (w/v, final concentration)  $\text{NaHCO}_3$ . The coupler mixture (50 mM HEPES, 400  $\mu\text{M}$  ATP, 10  $\mu\text{M}$   $\text{NAD}^+$ , 30  $\mu\text{M}$  CoA, 0.07 U/ $\mu\text{L}$  CS, 0.04 U/ $\mu\text{L}$  MDH, 50  $\mu\text{M}$  ACS, 100 mM NaCl, 3 mM KCl, 50 mM MgCl, 2.5 mM L-malic acid, pH 8) were incubated for 20 minutes at room temperature and added to each quenched reaction (at a ratio of 10  $\mu\text{L}$  coupler mix/65  $\mu\text{L}$  reaction) in a 96 well black plate. The reactions were incubated at room temperature for 40 minutes and the NADH fluorescence (ex. = 340 nm, em. = 460 nm) was measured.

Variation in data between experiments was <10 %. The steady state kinetic parameters for the Ac-KGGAKac-COO<sup>-</sup> peptide were determined from fitting the Michaelis-Menten equation to the concentration dependence of HDAC-catalyzed deacetylation. Substrate inhibition is observed for the peptides Ac-KGGAKac-NH<sub>2</sub> and Ac-KGGAKacW-NH<sub>2</sub>, therefore the kinetic parameters for these substrates were determined by fitting Equation 1 to the dependence of the initial velocities on peptide concentration. Equation 1 was derived from rearrangement of the Michaelis–Menten incorporating a hill constant (n) to report the value of  $k_{\text{cat}}/K_M$  and the standard error directly from the output.

$$\text{Equation 1: } \frac{v_0}{[E]} = \frac{\frac{k_{\text{cat}}}{K_M} \times [S]}{1 + \frac{[S]}{K_M} + \frac{[S]}{K_M} \times \frac{[S]^n}{K_I^n}}$$

### **Optimized continuous coupled acetate assay**

The 96 well plates were soaked (>3 hours) in 100 mM divalent metal-free EDTA to strip the plate of contaminating metal. The continuous assay buffer (50 mM HEPES, 400  $\mu$ M ATP, 10  $\mu$ M NAD<sup>+</sup>, 30  $\mu$ M CoA, 0.07 U/ $\mu$ L CS, 0.04 U/ $\mu$ L MDH, 50  $\mu$ M ACS, 127 mM NaCl, 2.7 mM KCl, 2.5 mM L-malic acid, pH 8) was incubated with Chelex resin for 1 hour at room temperature. The mixture was clarified by centrifugation at 16,800 x g for 2 minutes and the supernatant was collected. Then 6 mM magnesium was added to the buffer and the mixture was incubated for 20 minutes to allow NAD<sup>+</sup>/malate and NADH/OAA to equilibrate. The peptide (100  $\mu$ M final concentration Ac-KGGAKac-NH<sub>2</sub>) in HDAC8 assay buffer was added to this assay mixture at a ratio of 2:1, respectively. The reaction was initiated with the addition of Co(II)-HDAC8 (0.5 – 1  $\mu$ M final concentration) and deacetylation was measured from the time-dependent increase in NADH fluorescence (ex. = 340 nm, em. = 460 nm).

### **Fluorescamine assay**

The peptide substrate Ac-KGGAKac-COO<sup>-</sup> (0 - 1600  $\mu$ M) was preincubated in HDAC8 assay buffer (above) at 30°C for 10 minutes. The reactions were initiated by adding 0.5  $\mu$ M HDAC8, and quenched by addition of 1  $\mu$ M TSA after 0, 30, 60, and 90 minutes of incubation. The reactions were used immediately or flash frozen and stored at -20°C. Upon thawing, solutions were filtered through Pall 10K mwco NanosepMF Centrifugal devices to remove HDAC8. The flow-through (80  $\mu$ L) was mixed with 50  $\mu$ L of 1 M boric acid (pH 9) and the mixture was added to a Corning 96 well black plate. 33  $\mu$ L of 4.3 mM fluorescamine (dissolved in acetone) was then added to the sample, incubated at room temperature for 10 minutes, and fluorescence (ex. = 340

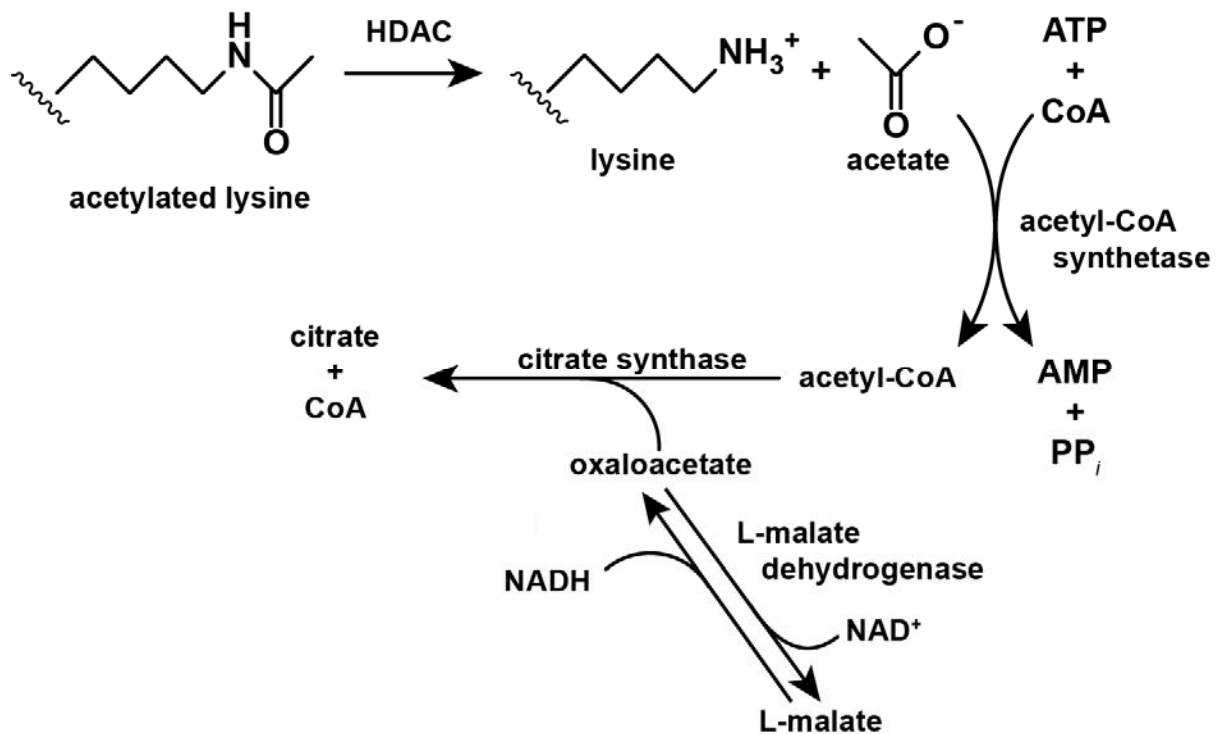


nm, em. = 460 nm) was measured[28]. A standard curve was created using N- $\alpha$ -acetyl lysine methyl ester (0 – 5  $\mu$ M). To measure the peptide concentration, peptides were diluted into 1 M borate, pH 9, and 0.56 mM fluorescamine was added. The mixture was incubated at room temperature for 10 min and the fluorescence (ex. = 340 nm, em. = 460 nm) was measured. The peptide concentration was determined from the standard curve.

## **Results**

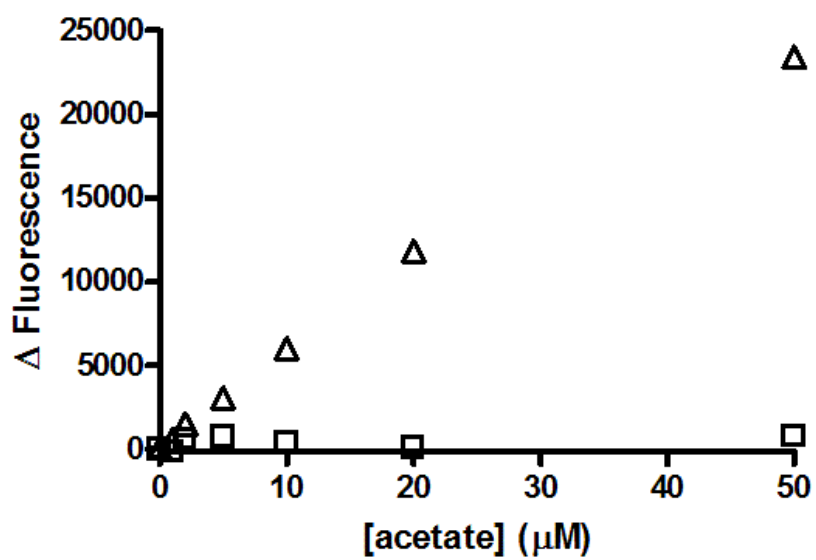
### **Assay**

An acetate detection assay was initially described in the Official Collection of Assays according to § 35 of German food law[29] and a kit containing the assay components is distributed by the R-biopharm company (Figure 2.1). This assay system couples enzymatic reactions that produce one molecule of citrate, CoA, and NADH per molecule of acetate. The NADH concentration is monitored using fluorescence and/or absorbance, allowing determination of the acetate concentration from a spectroscopic signal. This assay is optimized to measure millimolar concentrations of acetate with a detection limit of  $\sim$ 100  $\mu$ M (Figure 2.2 and 2.3), which is not sufficiently sensitive to measure the steady state kinetic parameters of many enzymes, including HDACs.



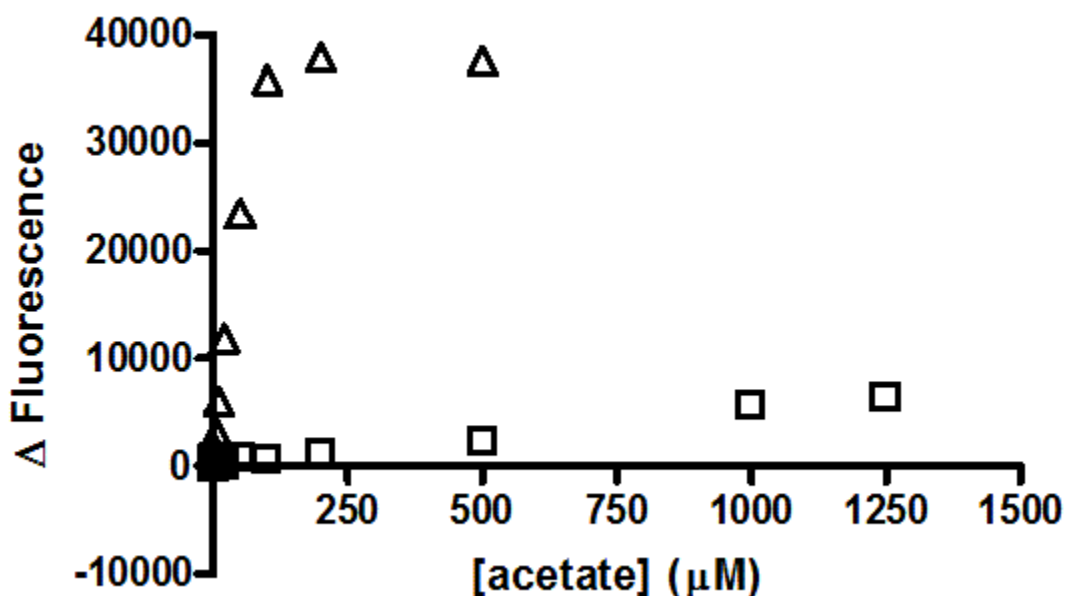
**Figure 2.1 HDAC Assay scheme**

Acetate is a product of deacetylation catalyzed by HDAC8. Acetate, ATP, and CoA are converted into acetyl-CoA, AMP, and inorganic pyrophosphate by acetyl-CoA synthetase. Acetyl-CoA and oxaloacetate are converted into citrate and CoA catalyzed by citrate synthase. Simultaneously, malate dehydrogenase catalyzes equilibration of NAD<sup>+</sup> and malate with NADH and oxaloacetate. When a molecule of oxaloacetate is removed from solution by formation of citrate, a molecule of NADH is formed. NADH concentrations are quantified using absorbance or fluorescence.



**Figure 2.2 Comparison of R-biopharm kit and optimized assay for detection of acetate**

The fluorescence change as a function of acetate concentration is measured using the R-biopharm kit (□) and the optimized acetate assay (Δ). The fluorescent signal is normalized to 0 μM acetate using the optimized assay. The signal from the optimized assay is much larger in the μM range, which is required to measure steady state turnover catalyzed by HDAC8. This signal is accurate to ~1 μM with a coefficient of variance equal to 2.2. The R-biopharm kit standard curve is linear at higher concentrations of acetate (250 – 1250 μM acetate) (Figure 2.3). Data from a single reaction is shown, the variation between experiments is <10 %.



**Figure 2.3 Extended acetate standard curve**

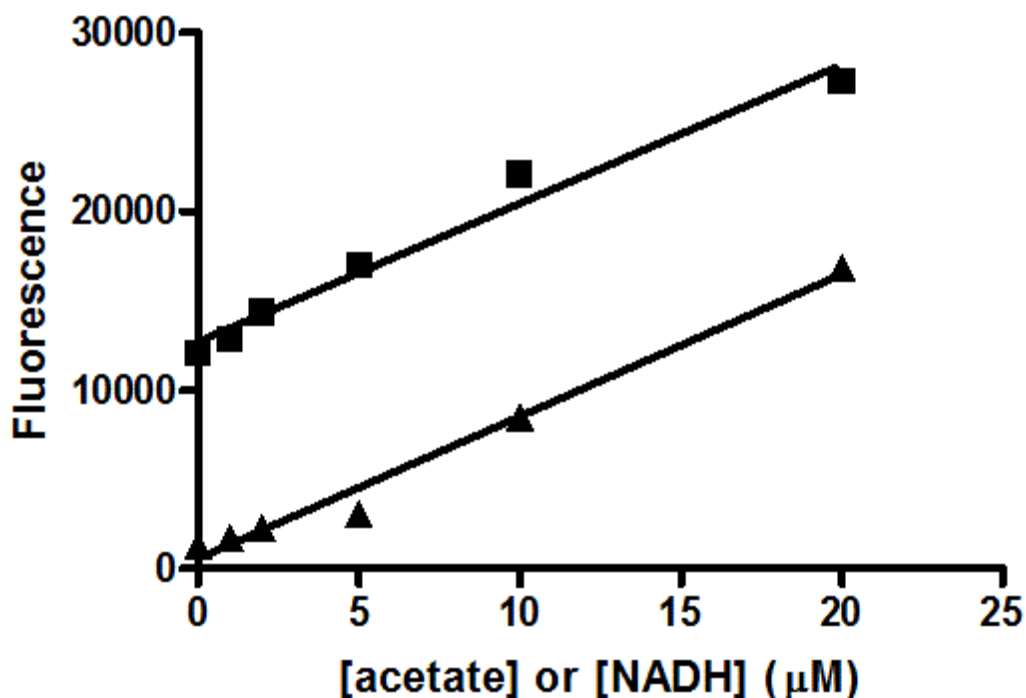
An acetate standard curve using the optimized coupled assay ( $\Delta$ ) and the kit ( $\square$ ) shows that the optimized coupled assay is linear from 0 – 50  $\mu\text{M}$ . The kit assay has a limit of detection of  $\sim 200$   $\mu\text{M}$  and is linear at higher concentrations. A 130:20 ratio of reaction to couplers was used for the optimized coupled assay while a ratio of 10:123 was used for the kit according to the instructions. Data from a single reaction is shown, the variation between experiments is  $<10\%$ .

### **Optimization of Assay for Measuring HDAC Activity**

The  $K_M$  values for HDAC-catalyzed hydrolysis of acetyl lysine residues in peptides are typically in the low to mid micromolar range. Therefore, to measure the initial rate ( $\leq 10\%$ ) of the reaction, the detection limit should be in the low micromolar range. To optimize the detection limit for the acetate-coupled assay, the signal to noise ratio was improved by: (1) using highly purified recombinant ACS, which decreased the background signal; (2) lowering the concentration of L-malic acid and  $\text{NAD}^+$  to decrease the background signal due to the

equilibrium formation of OAA and NADH; and (3) increasing the ratio of the sample to the coupling solution volume to improve the signal intensity. With these alterations, a linear standard curve from 0 to 50  $\mu\text{M}$  acetate with a limit of detection of  $\sim 1 \mu\text{M}$  (coefficient of variance 2.2; Figure 2.2) was produced using this assay. This assay can be altered to measure larger concentrations of acetate ( $>50 \mu\text{M}$ ) by adding higher concentrations of the limiting reagents CoA and  $\text{NAD}^+$ . This standard curve indicates that the optimized coupled assay is sensitive enough to measure the steady state kinetic parameters for HDACs and other enzymes.

Based on the assay design, each molecule of acetate should yield one molecule of NADH. To test this, the fluorescence change from the addition of acetate to the coupled assay was compared with the fluorescence of a comparable concentration of NADH (Figure 2.4). The slopes of the standard curves for NADH and acetate were  $770 \pm 61$  and  $790 \pm 50$  fluorescence units per  $\mu\text{M}$ , respectively. The equivalence of these slopes indicates that there is a one to one relationship between the concentration of acetate and the signal created by the production of NADH, allowing calculation of the acetate concentration from the fluorescence change.



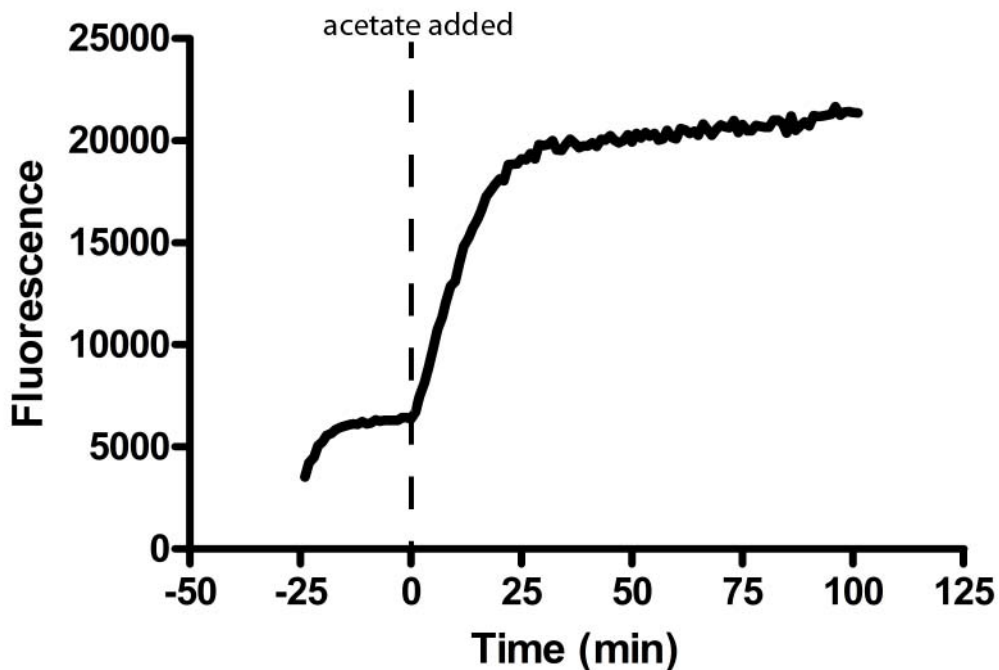
**Figure 2.4 Standard curve for the acetate assay compared to NADH fluorescence**

The fluorescence change observed upon addition of acetate to the coupled assay (■) is compared to the NADH fluorescence under comparable conditions (50 mM HEPES, 100 mM NaCl, 3 mM KCl, 50 mM MgCl, pH 8) (▲). The two slopes are equal, indicating that in the coupled assay one mole of NADH is formed per mole of acetate. The higher background observed in the acetate standard curve is a result of NADH formed before addition of acetate due to the equilibrium of the malate dehydrogenase-catalyzed reaction. Representative data is shown, the variation between experiments is <10 %.

### Stopped Assay

We first optimized the acetate assay in a stopped format to measure HDAC8 activity. After reacting HDAC8 with the substrate of interest, the reaction was quenched by the addition of HCl and flash frozen in liquid nitrogen. Upon thawing, the pH was neutralized by addition of NaHCO<sub>3</sub>. Using the Fluor de Lys assay to measure activity, the HCl solution quenches HDAC8

activity immediately (<10 sec), and HDAC8 activity is not restored upon neutralization (data not shown). The acetate concentration in this sample is then measured using the coupled assay. Under these optimized conditions, formation of NADH from the addition of acetate occurs within minutes. Upon mixing the coupling enzymes with the assay substrates, NAD<sup>+</sup> and L-malic acid equilibrate to form NADH and oxaloacetate. This equilibration is complete in 20 minutes (Figure 2.5), forming ~4 μM NADH, consistent with the equilibrium constant for the reaction catalyzed by malate dehydrogenase[30]. The reaction of up to 20 μM acetate is complete within 30 minutes and the signal remains stable for over an hour. The limiting step in this assay is the formation of citrate and CoA catalyzed by citrate synthase (CS). Therefore, the rate of acetate production can be increased by the addition of higher concentrations of CS, if needed.



**Figure 2.5 Time course for the formation of acetate**

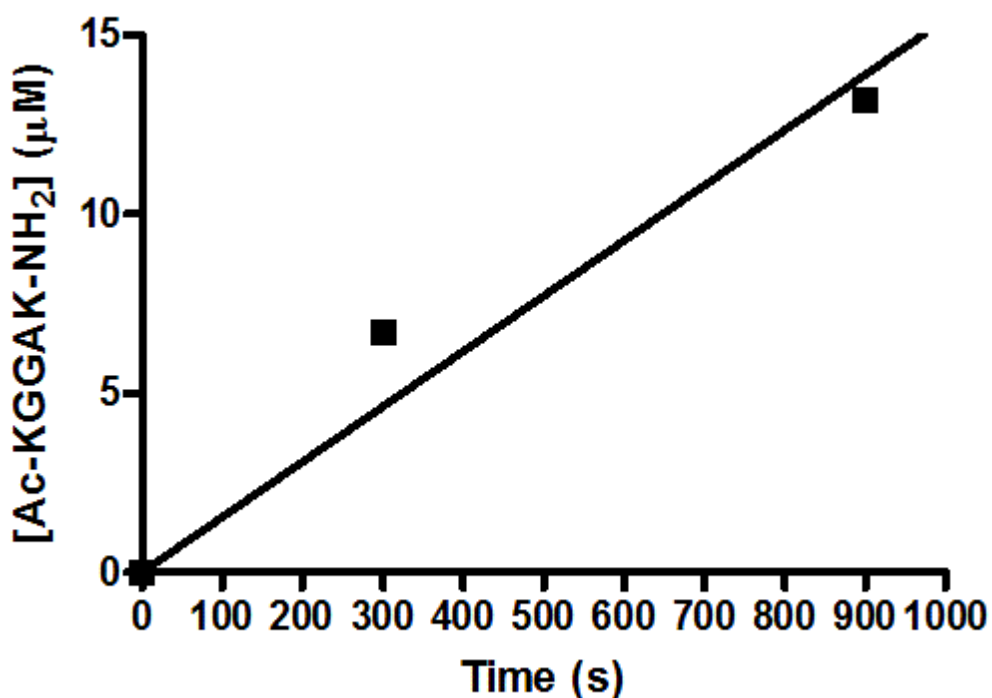
At -25 minutes the assay reagents are mixed and incubated at room temperature ( $\sim 25^{\circ}\text{C}$ ). Equilibration is complete in  $<20$  minutes, forming  $\sim 4 \mu\text{M}$  NADH. The reaction is initiated ( $t = 0$  min) by addition of  $20 \mu\text{M}$  acetate. The reaction reaches equilibrium at  $\sim 30$  min and the signal is stable for at least one hour.

To demonstrate the effectiveness of the acetate assay in measuring deacetylation, we compared the rate of HDAC8-catalyzed deacetylation determined using the coupled assay with the fluorescamine assay. Fluorescamine is a reagent that increases in fluorescence intensity upon reaction with primary amines[31] and therefore a fluorescent signal is coupled to the formation of lysine generated by HDAC-catalyzed deacetylation. We measured the reactivity of Co(II)-HDAC8 with an unlabeled peptide ( $\text{Ac-KGGAKac-COO}^-$ ) mimicking the H4 histone K16 acetylation site (H4 K16ac). HDAC8-catalyzed deacetylation of this peptide ( $200 \mu\text{M}$  peptide,



0.5  $\mu\text{M}$  HDAC8) measured by the coupled acetate assay and the fluorescamine assay yielded comparable rates within experimental error of  $0.0021 \pm 0.0003 \mu\text{M s}^{-1}$  and  $0.0027 \pm 0.0008 \mu\text{M s}^{-1}$ , respectively. Therefore, both assays measure the deacetylation rate and are viable for measuring HDAC8 activity, though the fluorescamine assay is less accurate for substrates containing multiple lysine side chains due to a higher signal to noise ratio.

To demonstrate that this optimized stopped assay can serve as a general acetate assay, we measured the initial rate of deacetylation of  $100 \mu\text{M}$  Ac-KGGAKac-NH<sub>2</sub> catalyzed by another HDAC isozyme,  $0.5 \mu\text{M}$  HDAC3/NCOR1 (Figure 2.6). The initial rate for this reaction was  $0.015 \pm 0.0016 \mu\text{M s}^{-1}$ , comparable to the value measured for HDAC8.



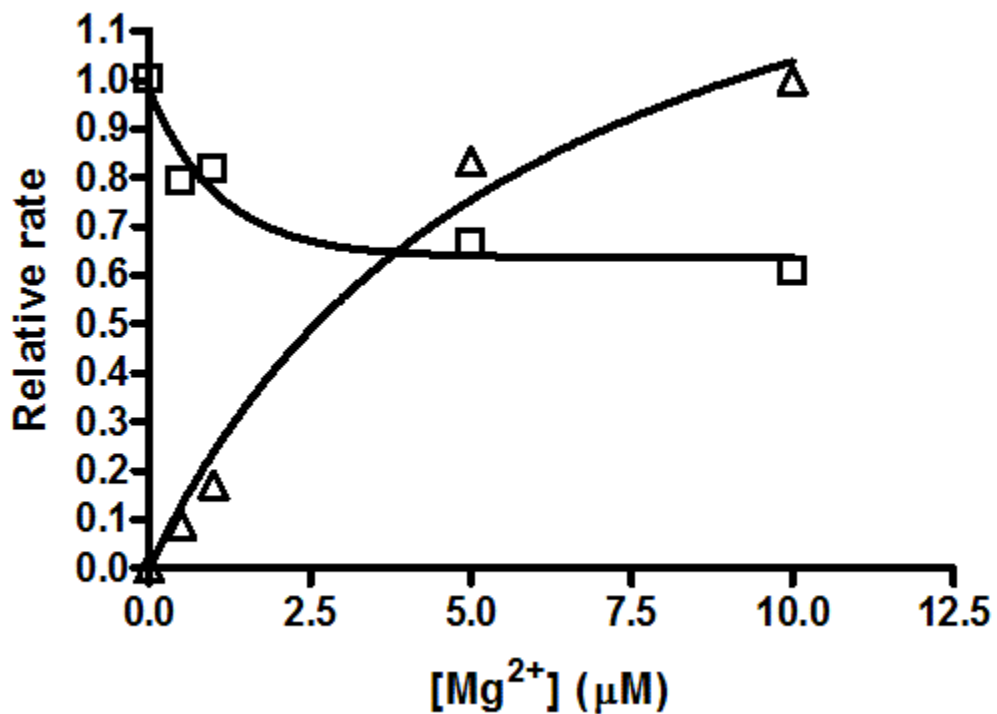
**Figure 2.6 HDAC3/NCOR1 stopped assay time course**

A time course measuring 0.5 µM HDAC3/NCOR1 deacetylation of 100 µM Ac-KGGAKac-NH<sub>2</sub>. An initial rate of  $0.022 \pm .0016 \mu\text{M s}^{-1}$  is measured. Data is linear within error over the 15 minute time course. Data from a single time course is shown and standard error was calculated from the fit to the three points.

### Continuous Assay

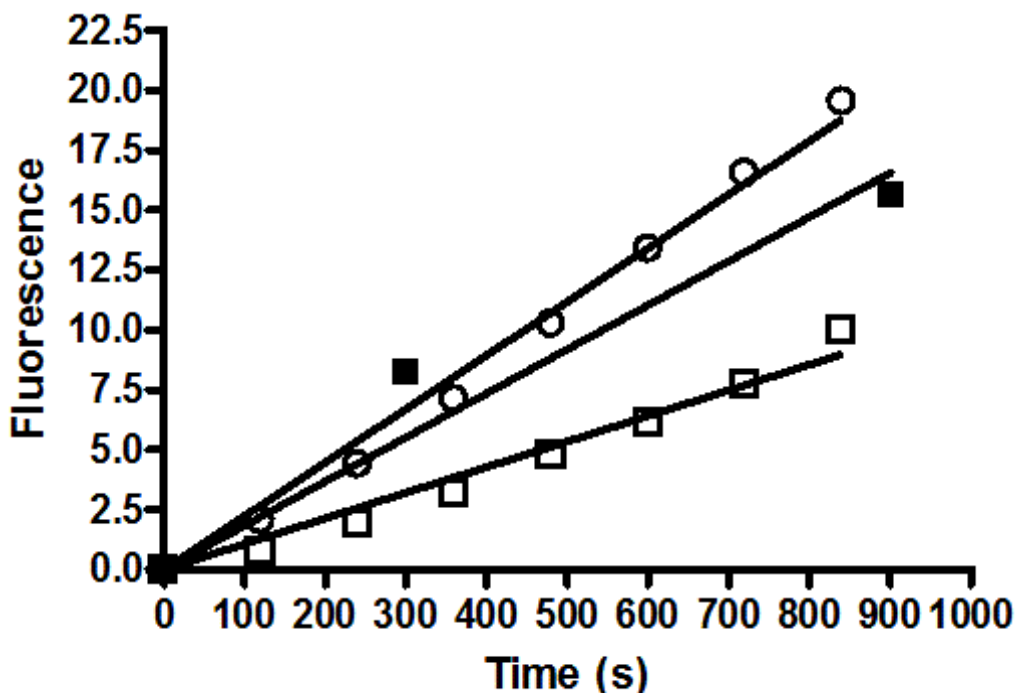
We next evaluated whether the coupled acetate assay could be carried out as a continuous, real-time assay to measure HDAC8 activity. Since HDAC8 is sensitive to inhibition by metals[9], and monovalent cations[15], the acetate coupling solutions were reformulated with concentrations of NaCl and KCl typically used to assay HDAC8 activity (127 mM and 2.7 mM respectively)[15] and treated with Chelex resin prior to addition of magnesium. The concentration of magnesium

was decreased to 2 mM to minimize inhibition of HDAC8 activity (~2-fold inhibition under these conditions) (Figure 2.7). To counteract the loss in activity of the coupling enzymes due to the lower concentration of magnesium, the concentration of these enzymes was increased by 2.3-fold to yield a final rate for the coupling reactions of  $0.046 \mu\text{M s}^{-1}$ . These assay conditions were used to measure HDAC8-catalyzed deacetylation of  $100 \mu\text{M}$  Ac-KGGAKac-NH<sub>2</sub> peptide yielding rates of  $0.018 \pm 0.00013 \mu\text{M s}^{-1}$  and  $0.028 \pm 0.00024 \mu\text{M s}^{-1}$  at  $0.5 \mu\text{M}$  and  $1 \mu\text{M}$  HDAC8, respectively (Figure 2.8). The linear dependence on the HDAC8 concentration demonstrates that the assay rate is not limited by the coupling reactions. Furthermore, the HDAC8 activity measured using the stopped assay ( $0.5 \mu\text{M}$  HDAC8 and  $100 \mu\text{M}$  Ac-KGGAKac-NH<sub>2</sub>) is  $0.033 \pm 0.0034 \mu\text{M s}^{-1}$  which is within the two-fold of the continuous assay measured rate, and represents the difference expected due to magnesium inhibition of Co(II)-HDAC8.



**Figure 2.7 Mg<sup>2+</sup> dependence of HDAC8 and acetate assay**

The value of  $k_{cat}/K_M$  for HDAC8 is decreased up to two-fold by millimolar concentrations of magnesium (□) (0.5 μM Co(II)-HDAC8, 50 μM Fluor de Lys substrate). The coupled assay reactions are increased up to 5-fold by similar concentrations of magnesium (Δ). A single representative curve is shown for each experiment.

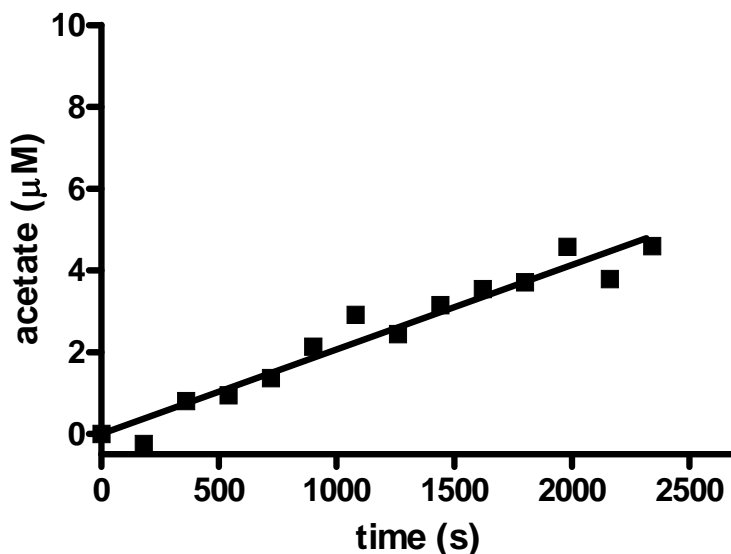


**Figure 2.8 Continuous assay to measure HDAC8 activity**

The initial rates for deacetylation of the Ac-KGGAKAc-NH<sub>2</sub> peptide (100 μM) measured using the acetate assay in a continuous formate are  $0.018 \pm 0.00013 \mu\text{M s}^{-1}$  and  $0.028 \pm 0.00024 \mu\text{M s}^{-1}$  for 0.5 μM (□) and 1 μM (○) Co(II)-HDAC8, respectively. A side-by-side acetate assay using the stopped formate (0.5 μM Co(II)-HDAC8, 100 μM Ac-KGGAKAc-NH<sub>2</sub>) has an initial rate of  $0.033 \pm 0.0021 \mu\text{M s}^{-1}$  (■). Error was calculated on the discrepancies between the data a line fit to that data from a single data set. The doubling of rate upon doubling of HDAC8 concentration shows that the coupling reactions are not rate limiting. Furthermore, as the continuous and stopped reaction rates are within error it suggests that both methods are feasible for the measurement of HDAC deacetylation.

Additionally, we measured the HDAC8-catalyzed deacetylation of H3/H4 tetramer acetylated using acetic anhydride, where ~100% of the lysines were acetylated as indicated by mass spectrometry (data not shown). The initial rate for deacetylation of ~0.076 μM acetylated H3/H4

tetramer catalyzed by 0.5  $\mu\text{M}$  Co(II)-HDAC8 is  $0.0021 \pm 0.0001 \mu\text{M s}^{-1}$  (Figure 2.9), demonstrating that this assay can measure deacetylation of both peptide and protein substrates.



**Figure 2.9 HDAC8-catalyzed deacetylation of pan-acetylated H3/H4 histone tetramer**

Histone tetramer was assembled from recombinantly expressed and purified histones H3 and H4 as in [32]. Tetramer was chemically acetylated using a method adapted from [33, 34]. In brief, tetramer was dialyzed in 2L of acetylation buffer (65.5 mM sodium bicarbonate pH 8.0, 2 M sodium chloride, 1 mM EDTA, 0.5 mM DTT). Dialyzed tetramer incubated on ice and 10  $\mu\text{L}$  of acetic anhydride was added to pan-acetylate the histones. 0.1 N NaOH was added dropwise, with stirring, to maintain pH 8. The reaction was incubated on ice for 40 minutes. The reaction was subsequently dialyzed in 2L acetylation buffer. Following dialysis, the tetramer was concentrated and exchanged into 50 mM HEPES pH 7.8, 300 mM NaCl, then incubated with Chelex resin at 4°C. The concentration of tetramer was measured using the BCA protein assay kit (Pierce). Tetramer was assayed using the continuous assay with 0.5  $\mu\text{M}$  Co(II)-HDAC8 and 0.076  $\mu\text{M}$  pan acetylated H3/H4 tetramer.

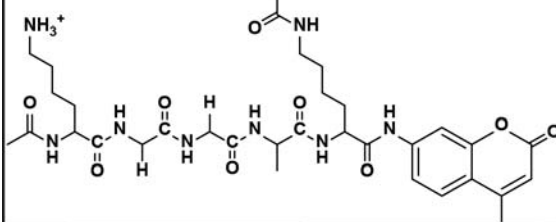
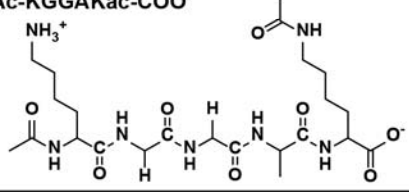
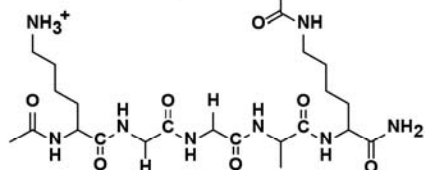
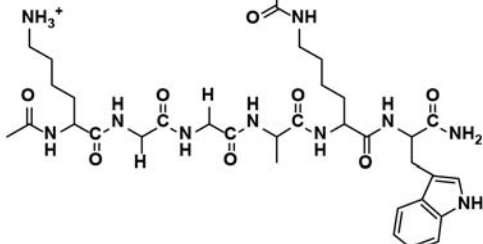
### Reactivity of HDAC8 with peptides

Using a mass spectrometric assay, Gurard-Levin *et al.*[23] previously demonstrated that HDAC8 catalyzes deacetylation of a peptide representing the local sequence of p53 containing a C-terminal methylcoumarin fluorophore significantly faster than a comparable peptide with a C-

terminal cysteine followed by an amide terminus. To further analyze the recognition of the peptide C-terminus by HDAC8, we measured Co(II)-HDAC8-catalyzed deacetylation using the stopped coupled assay of an H4 peptide mimic containing varied C-termini, including: a carboxylate (Ac-KGGAKac-COO<sup>-</sup>); an amide (Ac-KGGAKac-NH<sub>2</sub>); and a tryptophan capped by an amide (Ac-KGGAKacW-NH<sub>2</sub>) (Figure 2.10 and Table 2.1). The steady state kinetic parameters for Ac-KGGAKac-COO<sup>-</sup> were determined from fitting the Michaelis-Menten equation to the concentration dependence of activity, yielding values of  $k_{cat}/K_M$ ,  $k_{cat}$  and  $K_M$  of 56 M<sup>-1</sup> s<sup>-1</sup>, 0.041 s<sup>-1</sup> and 730 μM, respectively (Table 2.1). For the Ac-KGGAKac-NH<sub>2</sub> and Ac-KGGAKacW-NH<sub>2</sub> peptides, modest to substantial inhibition was observed at higher peptide concentrations. This inhibition is not due to metal contamination of the peptide[9] as pre-incubation of the peptide with Chelex 100 resin had no effect on the observed activity (Table 2.1). Furthermore, the pH of the peptides was measured (pH ~5) as a result, it is likely buffered HEPES used in the reactions and is also unlikely to affect the HDAC8 deacetylation rate. Therefore, these data indicate that the activity is inhibited by high substrate concentrations. As many HDAC isozymes form multi-protein complexes[35, 36], the peptides may inhibit HDAC8 activity by binding to non-active site protein-protein interaction sites. For these peptides the kinetic parameters were determined by fitting an equation including terms for substrate inhibition (Equation 1) to the data; the values of  $k_{cat}/K_M$  for deacetylation of Ac-KGGAKacW-NH<sub>2</sub> and Ac-KGGAKac-NH<sub>2</sub> are 1200 ± 250 M<sup>-1</sup>s<sup>-1</sup> and 980 ± 47 M<sup>-1</sup>s<sup>-1</sup>, respectively (Table 2.1). However, since the  $K_M$  and  $K_I$  parameters are coupled, independent values for these parameters could not be accurately determined. Substrate inhibition by the Ac-KGGAKac-NH<sub>2</sub> peptide appears to be cooperative with a Hill coefficient ( $n$ ) that is larger than 1. In general, substrate inhibition

significantly complicates the analysis of the reactivity of HDAC8 with peptide libraries, especially if activity is measured at a single peptide concentration.



Peptide	$k_{cat}/K_M$ ( $M^{-1} s^{-1}$ )	$k_{cat}$ ( $s^{-1}$ )	$K_M$ ( $\mu M^{-1}$ )
<b>Fluor de Lys H4 K16ac peptide</b> 	2800 <sup>a</sup> ± 200	0.90 <sup>a</sup> ± 0.03	320 <sup>a</sup> ± 20
<b>Non-fluorophore conjugated H4 K16ac peptide</b> <b>Ac-KGGAKac-COO<sup>-</sup></b> 	56 <sup>b</sup> ± 14	0.041 <sup>b</sup> ± 0.0042	730 <sup>b</sup> ± 160
<b>Non-fluorophore conjugated H4 K16ac peptide</b> <b>Ac-KGGAKac-NH<sub>2</sub></b> 	980 <sup>c</sup> ± 47	N.D. <sup>d</sup>	N.D. <sup>d</sup>
<b>Non-fluorophore conjugated H4 K16ac peptide</b> <b>Ac-KGGAKacW-NH<sub>2</sub></b> 	1200 <sup>c</sup> ± 250	N.D. <sup>d</sup>	N.D. <sup>d</sup>

**Table 2.1 Structure and kinetic constants for HDAC8-catalyzed deacetylation<sup>a</sup>**

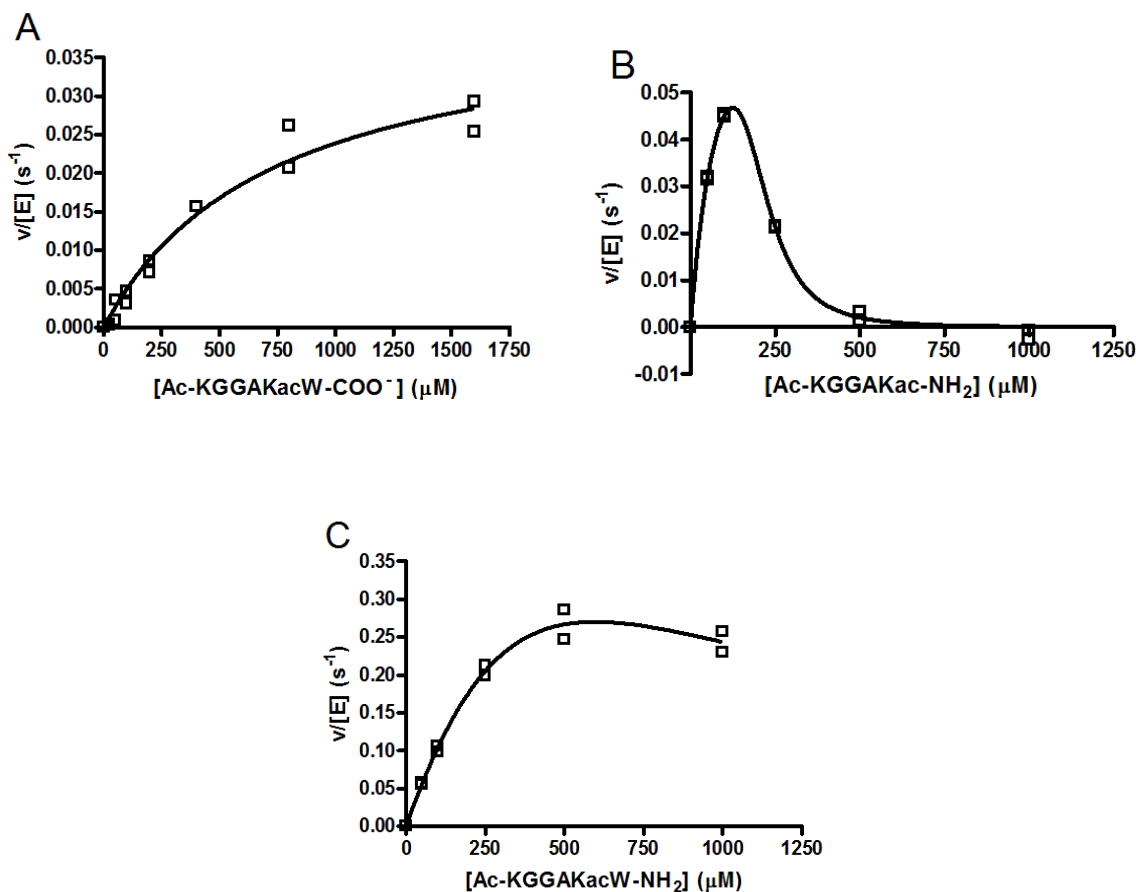
<sup>a</sup>The steady-state kinetic parameters and standard errors for HDAC8-catalyzed are measured as described in the legend of Figure 2.10.

<sup>b</sup>Values reported in ref [25].

<sup>c</sup>Kinetic parameters determined using Equation 1, including substrate inhibition.

<sup>d</sup>N.D. indicates values that were not determined as a result of substrate inhibition.

The four peptides differ in their C-terminus; the Fluor de Lys peptide contains a methylcoumarin fluorophore while the non-fluorophore conjugated H4 K16ac peptides contain a C-terminal carboxylate, carbamide, or tryptophan followed by a carbamide. The kinetic constants of these peptides suggest that the differences in the peptides affect both substrate binding and chemistry. Standard error is reported.



**Figure 2.10 Dependence of HDAC8-catalyzed deacetylation on the concentration of H4 K16ac peptide mimics**

The steady state kinetics of deacetylation catalyzed by HDAC8 were measured at 30°C in 50 mM HEPES, 137 mM NaCl, 2.7 mM KCl, pH 8, 0.85  $\mu M$  HDAC8. The solid line is a fit of the Michaelis-Menten to the initial rates for deacetylation of a peptide ending in A) carboxylic acid (Ac-KGGAKac-COO<sup>-</sup>). B) The peptide ending in an amide (Ac-KGGAKac-NH<sub>2</sub>). C) The peptide ending in a tryptophan and capped with an amide (Ac-KGGAKacW-NH<sub>2</sub>). Equation 1 was fit to the data for the Ac-KGGAKac-NH<sub>2</sub> and Ac-KGGAKacW-NH<sub>2</sub>. The solid lines shown are the best fits with the following values: Ac-KGGAKac-NH<sub>2</sub>,  $k_{cat}/K_M$ ,  $K_M$ ,  $K_I$ , and  $n$  values of  $980 \pm 47 M^{-1} s^{-1}$ ,  $90 \pm 21 \mu M$ ,  $190 \pm 15 \mu M$ , and  $3.9 \pm 0.6$ , respectively and Ac-KGGAKacW-NH<sub>2</sub>,  $k_{cat}/K_M$ ,  $K_M$ , and  $K_I$  values of  $1250 \pm 250 M^{-1} s^{-1}$ ,  $890 \pm 2100 \mu M$ , and  $440 \pm 1600 \mu M$  respectively. As the  $K_M$  and  $K_I$  values are highly correlated, these values are not accurate.

These data demonstrate that the structure of the C-terminal moiety significantly affects the reactivity of Co(II)-HDAC8 with peptide substrates. Furthermore, a comparison of these data to the  $k_{cat}/K_M$  value of  $2800 \text{ M}^{-1}\text{s}^{-1}$ [25] for HDAC8-catalyzed deacetylation of the Ac-KGGAKac-methylcoumarin peptide demonstrates that the methylcoumarin fluorophore enhances (up to 50-fold) the catalytic efficiency of HDAC8-catalyzed deacetylation (Table 2.1).

## **Discussion**

The current HDAC assays have limitations that do not allow in-depth and/or high throughput analysis of HDAC substrate specificity. The first kinetic measurements of HDAC activity employed radioactively labeled peptides and proteins and detected the formation of radiolabeled acetate[19, 20]. While this assay is effective and sensitive, radioactive peptide substrates are expensive to produce, and many radiolabeled protein substrates are acetylated non-specifically, preventing the determination of detailed kinetics at specific sites. Recently, the Fluor de Lys assay (Biomol) has become popular for measuring HDAC kinetics[9, 11-17]. This assay uses a methylcoumarin fluorophore-conjugated peptide, which upon deacetylation becomes a substrate for the serine protease trypsin, cleaving the fluorophore and altering the fluorescence spectrum of the methylcoumarin. The rate of deacetylation is measured by the change in the fluorescence signal[12]. However, the methylcoumarin fluorophore can interact with HDACs and alter the kinetics of deacetylation[23] (as discussed later). Thus, the results obtained using this method may not accurately report the selectivity of HDACs for native peptides and proteins.

Additionally, because the fluorophore is located immediately on the C-terminal side of the acetyl lysine moiety, this method cannot be used to determine the preference for sequences downstream

of the acetylated lysine. Many of these limitations were solved by a mass spectrometric assay developed by the Mrksich group; this assay uses MALDI mass spectrometry of peptides attached to a gold surface to observe the mass difference caused by deacetylation[23]. This assay is effective for peptide substrates and can be carried out in a high throughput manner, however a MALDI mass spectrometer is required. Additionally, this method requires the inclusion of a cysteine in the peptide sequence to conjugate the peptide to the plate via a maleimide linkage. HPLC separation of acetylated and deacetylated peptides on a C18 column is another method employed to determine deacetylation kinetics[18]. This assay quantifies product formation by absorbance (230 or 280 nm) and can measure the deacetylation of any peptide, yet it suffers from labor intensive techniques such as the determination of peak elution times making it not ideal for a high throughput format. Furthermore, it is not easily transitioned to assaying the deacetylation of proteins due to the difficulty of separating full length proteins differing by a single acetylated lysine. The assay presented in this paper provides an alternative that overcomes many of the limitations of the previous assays. This acetate assay is versatile; either fluorescence or absorbance can be measured using a cuvette or 96-well plate and in a stopped or continuous format. The continuous method is well-suited for high-throughput screening. Peptide substrates for the assay are inexpensive to purchase as they do not require conjugated fluorophores, and results can be quantified using an acetic acid standard curve. Furthermore, the assay can measure deacetylation of proteins (Figure 2.9). However, the acetate assay cannot be used with fluorescent peptides (i.e. methylcoumarin-conjugated peptides) that absorb and emit at wavelengths similar to NADH, nor can it be used on cell lysate extracts containing NADH, other metabolites, or various other enzymes. Overall, this assay provides a stable and sensitive platform for the measurement of deacetylation with few limitations. Finally, as this assay has

been optimized to measure micromolar acetate concentrations, it can be used to assay other enzymes that produce acetate as a product and have  $K_M$  values in the micromolar range, including other HDAC isozymes (Figure 2.6).

While most kinetic measurements of HDAC8 have been performed using the Fluor de Lys assay[9, 11, 13-15], the Mrksich group used a mass spectrometric assay to demonstrate that a peptide mimicking the p53 transcription factor was deacetylated significantly slower than the same peptide containing a methyl-coumarin fluorophore[23]. However, the steady state kinetic parameters were not determined for these substrates. Using the optimized coupled assay that measures the formation of acetate, we demonstrated that substitution of the coumarin fluorophore with a carboxylate lowered the value of  $k_{cat}/K_M$  for Co(II)-HDAC8 by 50-fold, resulting from a 22-fold reduction in the value of  $k_{cat}$  and a 2.3-fold increase in the  $K_M$  value (Table 2.1). As the value of  $K_M$  is relatively high (730  $\mu\text{M}$ ) and the value of  $k_{cat}$  is relatively low (0.041  $\text{s}^{-1}$ ) compared to other enzymes acting under diffusion control[37], it is likely that substrate dissociation is faster than deacetylation, indicative of a rapid equilibrium substrate binding model. This assumption is further validated by the increase in  $k_{cat}$  measured for trifluoroacetyl lysine substrates[18, 38], suggesting that deacetylation is the rate-limiting step for  $k_{cat}$ . This conclusion is further bolstered by the low  $k_{cat}/K_M$  values for HDAC8 deacetylation of peptide substrates ( $<10^4$ ) further suggesting that chemistry is rate limiting. Based on this assumption,  $K_M$  reflects  $K_D$  for the peptide. Therefore,  $\Delta\Delta G_{\text{binding}}$ , calculated from the alteration in the  $K_M$  values<sup>3</sup>, indicative of the additional binding affinity conferred by the methylcoumarin fluorophore relative to a carboxylate at the C-terminus, is equal to  $\sim 0.45$  kcal/mol. This alteration in binding energy is modest but within the range of energy due to the addition of a

---

<sup>3</sup>  $\Delta\Delta G_{\text{binding}} = RT \ln(K_M \text{ Fluor de Lys peptide}) - RT \ln(K_M \text{ non-fluorophore conjugated peptide})$

single pi-pi interaction[39, 40] or hydrogen bond[37]. Crystal structures of a methylcoumarin-conjugated peptide bound to HDAC8 visualize interactions between the methylcoumarin and the side chain of Tyr100[13, 41]. These structures suggest that the enhanced binding energy for the methylcoumarin peptide results from a combination of interactions between the C-terminus of the peptide and the hydrophobic cavity formed by the L1, L7, and L8 loops, and interactions between the aromatic C-terminal residue of the peptide and Tyr100 on the L2 loop of HDAC8.

The steady state kinetic parameters for deacetylation catalyzed by HDAC8 demonstrate that the significant (50-fold) enhancement in Co(II)-HDAC8  $k_{cat}/K_M$  for the Ac-KGGAKac-methylcoumarin peptide compared to the Ac-KGGAK-COO<sup>-</sup> peptides is largely due to an increase in the  $k_{cat}$  value. The change in the stabilization of the transition state relative to the unbound ground state ( $\Delta\Delta G^\ddagger$ )<sup>4</sup> of 2.3 kcal/mol likely results from a combination of altered electrostatic and pi-pi interactions between the peptide and HDAC8 that enhance optimal positioning of the peptide and side chains in the active site to efficiently catalyze deacetylation. The effects of the C-terminal interactions of the peptide with Co(II)-HDAC8 on  $k_{cat}$  and  $K_M$  are consistent with data demonstrating that L2 loop residues are important for both binding and catalysis; mutations at Asp101 in HDAC8 lead to both higher  $K_M$  and lower  $k_{cat}$  values[13], compared to the wildtype enzyme. However, these mutations do not lead to observable alterations in the crystal structure of inhibitor-bound HDAC8[13, 41], suggesting that the activity decrease may be due to an alteration in the HDAC8 dynamics. The catalytic efficiency ( $k_{cat}/K_M$ ) of Co(II)-HDAC8 with the Ac-KGGAKac-NH<sub>2</sub> peptide is enhanced 9-fold ( $\Delta\Delta G^\ddagger = 1.6$

---

<sup>4</sup>  $\Delta\Delta G^\ddagger (k_{cat}/K_M) = RT \ln(k_{cat}/K_M_{peptide1}) - RT \ln(k_{cat}/K_M_{peptide2})$

kcal/mol)<sup>5</sup> compared to the Ac-KGGAKac-COO<sup>-</sup> peptide indicating that electrostatic interactions between the C-terminus and HDAC8 impair peptide binding and/or reactivity. Addition of a tryptophan (Ac-KGGAKacW-NH<sub>2</sub>) or methylcoumarin moiety to the peptide increases net transition state stabilization by 0.1 – 0.5 kcal/mol, providing an estimate of the enhancement of the catalytic efficiency by base stacking with Tyr100. Previous studies performed by the Mrksich lab show that HDAC8 catalyzes deacetylation of peptides containing a phenylalanine on the C-terminal side of the acetyl lysine faster than substrates containing any other amino acid, including tryptophan, at that position[21, 23]. These data suggest that both amino acid hydrophobicity and volume play a role in substrate preference. However, a direct comparison of the reactivity of peptides in this paper compared to those in Gurard-Levin *et al.*[21, 23] is complicated by differences in peptide length, the sequence at other positions of the peptide, and method of kinetic measurement. Further kinetic experiments utilizing the Ac-KGGAKacW-NH<sub>2</sub> peptide will be necessary to determine exactly how the studies correlate to each other.

Interestingly, Co(II)-HDAC8-catalyzed deacetylation of acetyl lysine peptides have low values for  $k_{cat}/K_M$  (56 – 2800 M<sup>-1</sup> s<sup>-1</sup>) compared to enzymes that are limited by diffusion[37], suggesting either that these peptides are poor substrates for this enzyme or that the low activity is biologically relevant for control of enzyme activity[42]. The *in vivo* catalytic efficiency could be enhanced either by additional interactions with the protein substrates or by additional cofactors or binding partners. Furthermore, the measured rate constants for Co(II)-HDAC8-catalyzed deacetylation of Ac-KGGAKac-COO<sup>-</sup> and high concentrations of Ac-KGGAKac-NH<sub>2</sub> peptides are similar to the rates measured for deacetylation catalyzed by various class II HDAC isozymes

---

<sup>5</sup>  $\Delta\Delta G_{binding} = RT \ln(K_M \text{ Fluor de Lys peptide}) - RT \ln(K_M \text{ non-fluorophore conjugated peptide})$

(HDAC7, 10)[17, 43] suggesting that the low activity of these isozymes in the Fluor de Lys assay may reflect decreased enhancement of reactivity by aromatic C-terminal moieties.

## **Acknowledgements**

We thank Dr. Eric Drake and Dr. Andrew Gulick for their generous contribution of the ACS plasmid and subsequent help in ACS purification. We thank Byungchul Kim for his help in using 96 well plates for this assay. We also thank members of the Fierke lab for their insight, discussion, and critiques of the material within the paper. This material was supported, in whole or in part, by the National Institutes of Health Grants, NIGMS GM40602 (CAF), T32-GM-008353 (NAW), and T32-GM-008597 (CAP), and the National Science Foundation Graduate Fellowship DGE0718128 (NAW).

## **Bibliography**

- [1] S. Khochbin, A. Verdel, C. Lemercier, D. Seigneurin-Berny, Functional significance of histone deacetylase diversity, *Curr Opin Genet Dev*, 11 (2001) 162-166.
- [2] X.J. Yang, E. Seto, The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men, *Nat Rev Mol Cell Biol*, 9 (2008) 206-218.
- [3] G.A. Khoury, R.C. Baliban, C.A. Floudas, Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database, *Scientific reports*, 1 (2011).
- [4] M.A. Glozak, N. Sengupta, X. Zhang, E. Seto, Acetylation and deacetylation of non-histone proteins, *Gene*, 363 (2005) 15-23.
- [5] C. Choudhary, C. Kumar, F. Gnad, M.L. Nielsen, M. Rehman, T.C. Walther, J.V. Olsen, M. Mann, Lysine acetylation targets protein complexes and co-regulates major cellular functions, *Science*, 325 (2009) 834-840.



- [6] S. Zhao, W. Xu, W. Jiang, W. Yu, Y. Lin, T. Zhang, J. Yao, L. Zhou, Y. Zeng, H. Li, Y. Li, J. Shi, W. An, S.M. Hancock, F. He, L. Qin, J. Chin, P. Yang, X. Chen, Q. Lei, Y. Xiong, K.L. Guan, Regulation of cellular metabolism by protein lysine acetylation, *Science*, 327 (2010) 1000-1004.
- [7] P. Marks, R.A. Rifkind, V.M. Richon, R. Breslow, T. Miller, W.K. Kelly, Histone deacetylases and cancer: causes and therapies, *Nature reviews. Cancer*, 1 (2001) 194-202.
- [8] M. Lemoine, A. Younes, Histone deacetylase inhibitors in the treatment of lymphoma, *Discov Med*, 10 (2010) 462-470.
- [9] S.L. Gantt, S.G. Gattis, C.A. Fierke, Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion, *Biochemistry*, 45 (2006) 6170-6178.
- [10] L. Tong, J.M. Denu, Function and metabolism of sirtuin metabolite O-acetyl-ADP-ribose, *Biochim Biophys Acta*, 1804 (2010) 1617-1625.
- [11] D. Riester, C. Hildmann, S. Grunewald, T. Beckers, A. Schwienhorst, Factors affecting the substrate specificity of histone deacetylases, *Biochem Biophys Res Commun*, 357 (2007) 439-445.
- [12] D. Wegener, F. Wirsching, D. Riester, A. Schwienhorst, A fluorogenic histone deacetylase assay well suited for high-throughput activity screening, *Chemistry & biology*, 10 (2003) 61-68.
- [13] D.P. Dowling, S.L. Gantt, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structural studies of human histone deacetylase 8 and its site-specific variants complexed with substrate and inhibitors, *Biochemistry*, 47 (2008) 13554-13563.
- [14] D.P. Dowling, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structures of metal-substituted human histone deacetylase 8 provide mechanistic inferences on biological function, *Biochemistry*, 49 (2010) 5048-5056.
- [15] S.L. Gantt, C.G. Joseph, C.A. Fierke, Activation and inhibition of histone deacetylase 8 by monovalent cations, *J Biol Chem*, 285 (2010) 6036-6043.
- [16] S. Haider, C.G. Joseph, S. Neidle, C.A. Fierke, M.J. Fuchter, On the function of the internal cavity of histone deacetylase protein 8: R37 is a crucial residue for catalysis, *Bioorg Med Chem Lett*, 21 (2011) 2129-2132.
- [17] B.E. Schultz, S. Misialek, J. Wu, J. Tang, M.T. Conn, R. Tahilramani, L. Wong, Kinetics and comparative reactivity of human class I and class IIb histone deacetylases, *Biochemistry*, 43 (2004) 11083-11091.

- [18] B.C. Smith, J.M. Denu, Acetyl-lysine analog peptides as mechanistic probes of protein deacetylases, *J Biol Chem*, 282 (2007) 37256-37265.
- [19] J.J. Buggy, M.L. Sideris, P. Mak, D.D. Lorimer, B. McIntosh, J.M. Clark, Cloning and characterization of a novel human histone deacetylase, HDAC8, *Biochem J*, 350 Pt 1 (2000) 199-205.
- [20] I. Van den Wyngaert, W. de Vries, A. Kremer, J. Neefs, P. Verhasselt, W.H. Luyten, S.U. Kass, Cloning and characterization of human histone deacetylase 8, *FEBS Lett*, 478 (2000) 77-83.
- [21] Z.A. Gurard-Levin, K.A. Kilian, J. Kim, K. Bahr, M. Mrksich, Peptide arrays identify isoform-selective substrates for profiling endogenous lysine deacetylase activity, *ACS Chem Biol*, 5 (2010) 863-873.
- [22] Z.A. Gurard-Levin, M. Mrksich, The activity of HDAC8 depends on local and distal sequences of its peptide substrates, *Biochemistry*, 47 (2008) 6242-6250.
- [23] Z.A. Gurard-Levin, J. Kim, M. Mrksich, Combining mass spectrometry and peptide arrays to profile the specificities of histone deacetylases, *Chembiochem*, 10 (2009) 2159-2161.
- [24] A.S. Reger, J.M. Carney, A.M. Gulick, Biochemical and crystallographic analysis of substrate binding and conformational changes in acetyl-CoA synthetase, *Biochemistry*, 46 (2007) 6536-6546.
- [25] S.L. Gantt, Human Histone Deacetylase 8: Metal Dependence and Catalytic Mechanism, Chemistry, The University of Michigan, Ann Arbor, 2006.
- [26] J.E. Tropea, S. Cherry, D.S. Waugh, Expression and purification of soluble His(6)-tagged TEV protease, *Methods Mol Biol*, 498 (2009) 297-307.
- [27] S.C. Gill, P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, *Anal Biochem*, 182 (1989) 319-326.
- [28] X. Huang, M. Hernick, A fluorescence-based assay for measuring N-acetyl-1-D-myoinositol-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase activity, *Anal Biochem*, 414 (2011) 278-281.
- [29] H.U. Bergmeyer, J. Bergmeyer, M. Grassl, *Methods of enzymatic analysis*, 3rd ed., Verlag Chemie, Weinheim ; Deerfield Beach, Fla., 1983.

- [30] R.W. Guynn, H.J. Gelberg, R.L. Veech, Equilibrium constants of the malate dehydrogenase, citrate synthase, citrate lyase, and acetyl coenzyme A hydrolysis reactions under physiological conditions, *J Biol Chem*, 248 (1973) 6957-6965.
- [31] S. Udenfriend, S. Stein, P. Bohlen, W. Dairman, W. Leimgruber, M. Weigele, Fluorescamine: a reagent for assay of amino acids, peptides, proteins, and primary amines in the picomole range, *Science*, 178 (1972) 871-872.
- [32] K. Luger, A.W. Mader, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature*, 389 (1997) 251-260.
- [33] P. Bheda, S. Swatkoski, K.L. Fiedler, J.D. Boeke, R.J. Cotter, C. Wolberger, Biotinylation of lysine method identifies acetylated histone H3 lysine 79 in *Saccharomyces cerevisiae* as a substrate for Sir2, *Proceedings of the National Academy of Sciences of the United States of America*, 109 (2012) E916-925.
- [34] K. Marushige, Activation of chromatin by acetylation of histone side chains, *Proceedings of the National Academy of Sciences of the United States of America*, 73 (1976) 3937-3941.
- [35] E. Verdin, *Histone deacetylases : transcriptional regulation and other cellular functions*, Humana Press, Totowa, N.J., 2006.
- [36] N. Sengupta, E. Seto, Regulation of histone deacetylase activities, *J Cell Biochem*, 93 (2004) 57-67.
- [37] A. Fersht, *Structure and mechanism in protein science : a guide to enzyme catalysis and protein folding*, W.H. Freeman, New York, 1999.
- [38] D. Riester, D. Wegener, C. Hildmann, A. Schwienhorst, Members of the histone deacetylase superfamily differ in substrate specificity towards small synthetic substrates, *Biochem Biophys Res Commun*, 324 (2004) 1116-1123.
- [39] S. Tsuzuki, K. Honda, T. Uchimar, M. Mikami, K. Tanabe, Origin of attraction and directionality of the  $\pi/\pi$  interaction: model chemistry calculations of benzene dimer interaction, *J Am Chem Soc*, 124 (2002) 104-112.
- [40] M.O. Sinnokrot, E.F. Valeev, C.D. Sherrill, Estimates of the ab initio limit for  $\pi/\pi$  interactions: the benzene dimer, *J Am Chem Soc*, 124 (2002) 10887-10893.
- [41] A. Vannini, C. Volpari, P. Gallinari, P. Jones, M. Mattu, A. Carfi, R. De Francesco, C. Steinkuhler, S. Di Marco, Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex, *EMBO Rep*, 8 (2007) 879-884.

[42] C.A. Fierke, R.D. Kuchta, K.A. Johnson, S.J. Benkovic, Implications for enzymic catalysis from free-energy reaction coordinate profiles, *Cold Spring Harb Symp Quant Biol*, 52 (1987) 631-638.

[43] A. Schuetz, J. Min, A. Allali-Hassani, M. Schapira, M. Shuen, P. Loppnau, R. Mazitschek, N.P. Kwiatkowski, T.A. Lewis, R.L. Maglathin, T.H. McLean, A. Bochkarev, A.N. Plotnikov, M. Vedadi, C.H. Arrowsmith, Human HDAC7 harbors a class IIa histone deacetylase-specific zinc binding motif and cryptic deacetylase activity, *J Biol Chem*, 283 (2008) 11355-11363.

## Chapter 3

### Profiling small peptide substrates to identify HDAC8 recognition motif(s)<sup>10</sup>

#### Introduction

Multiple studies have utilized short peptides to determine HDAC substrate specificity[1-5]. These studies measure HDAC catalyzed deacetylation of short peptides with one or two amino acid changes. While these studies have been effective in identifying amino acid preferences in peptide substrates, they have been unable to identify *in vivo* HDAC substrates. One possibility for the lack of biological relevance is that combinations of three or more amino acids may be necessary for HDAC protein recognition. Alternatively, distal interactions not represented in the short peptides, may be necessary for protein recognition.

Of the metal dependent HDAC isozymes, HDAC8 has been historically regarded as the HDAC least likely to form a large protein complex that specifies substrate preference[6]. As a result, identifying HDAC8 specificity may be more straight forwards than for other HDAC isozymes. To determine the substrate preference of HDAC8, we utilized a computational approach which was informed by, and verified using experimental measurement of HDAC8 activity kinetics. In this chapter we developed an algorithm utilizing data from systematically altered peptide arrays and identified 26 corresponding to acetyl lysines in proteins which were grouped computationally into “good,” “bad,” and “mediocre” peptides. These peptides were subsequently synthesized and assayed, and based on the data, the algorithm was adjusted, and a new set of

---

<sup>10</sup> I performed all of the experimental studies with the peptide substrates (Figure 3.4, 3.7 and Table 3.4, 3.5) and I wrote all of the text pertaining to these experiments. Development of the computational algorithm was performed by Nawsad Alam and Lior Zimmerman.

peptides was predicted. One of the peptides in this second set is the most efficient peptide substrate for Zn(II)-HDAC8 measured to date. Overall, these results suggest that kinetic data of how one or two amino acid positions affect specificity is not sufficient to determine the substrate specificity of HDAC8, but a statistical compilation of ~6 amino acids surrounding the acetyl lysine provides a more accurate model for identifying efficient HDAC8 peptide substrates.

## **Materials and Methods**

### **Reagents**

ATP, Coenzyme A, NAD<sup>+</sup>, L-malic acid, citrate synthase, and malate dehydrogenase were purchased from Sigma. The initial peptide set was purchased from Sigma-Aldrich (~70 % pure). The second round of peptides was purchased from Peptide2.0 (~70 % pure). Cobalt and zinc were purchased as ICP standards from GFS Chemicals and the acetic acid standard was purchased from Ricca Chemical Company. Chelex 100 resin was purchased from Bio-Rad. All other materials were purchased from Fisher and were of a purity >95 % unless otherwise noted.

### **HDAC8 expression and assay**

HDAC8 was expressed and purified as previously described [7]. The peptides purchased from Sigma-Aldrich were solubilized in 50 % acetonitrile. The peptides purchased from Peptide 2.0 were initially solubilized in water, 50 % acetonitrile, or 10 % DMSO as needed. Once solubilized, the metals were removed from the peptides from Peptide 2.0 by incubation with ~30  $\mu$ L of Chelex resin per <5 mg peptide at 4°C for at least 16 hours. Peptides were quantified as previously described in Chapter 2. HDAC8 assays were performed as stopped assays as described in Chapter 2. Briefly, 1  $\mu$ M metal Zn(II)- or Co(II)-HDAC8 was added to peptides (concentrations specificities in the text) under standard assay conditions (2.7 mM KCl, 137 mM NaCl, 50 mM HEPES, pH 7.8) at 30°C. It should be noted that under the assay conditions (150  $\mu$ M substrate, 1  $\mu$ M HDAC8) the zinc concentration was 2.5  $\mu$ M sufficient to inhibit HDAC8[7]. As a result, the peptides were screened using Co(II)-HDAC8, which should only be partially

inhibited (~20%) by the presence of 1.5:1 zinc:HDAC8, instead of the Zn(II) enzyme which would be closer to ~70% inhibited. At various time points, the reaction was quenched by addition of 0.37% HCl. The reactions were neutralized by addition of 0.6% NaHCO<sub>3</sub>, were added to a coupling solution, and incubated at room temperature for 45 minutes. The increase in NADH reflected acetate production and was determined by fluorescence (ex. = 340 nm, em. = 460 nm). Linear portions of the time versus fluorescence curves were fit to a line were graphed versus HDAC8 concentration. The Michaelis-Menten equation was fit to these data.

### **Improving the assay conditions**

Prior to validation of the second model, we improved the assay conditions. In order to make the results more biologically relevant, changes were made to the assay. First, a method for chelating peptides was developed whereby Chelex 100 was added to each peptide and incubated overnight. Inhibitory concentrations of contaminating Zn<sup>2+</sup> were likely a result of the synthesis and solubilization process. Chelation allows for the use of Zn(II)-HDAC8, as inhibitory concentrations of zinc are no longer present in the assay. Zn(II)-HDAC8 is a more biologically relevant form of the HDAC8 enzyme when compared to the cobalt form. This allows us to better probe biologically relevant reactions than previously.

The first validation set of peptides contained a C-terminal carboxylate. Previous results have shown HDAC8 catalyzes the deacetylation of a peptide with a C-terminal carboxylate (immediately C-terminal of the acetyl lysine) 17.5-fold slower than the same peptide with a carboxylamine in the same position (Chapter 2). To determine if this is also the case when the C-terminal carboxylic acid is contained 3 amino acids C-terminal of the acetyl lysine, two peptides representing the Lys350 position of KAT6A were assayed. HDAC8 was incubated with two peptides (0 – 100 μM), one containing a C-terminal carbamide and the other a C-terminal carboxylic acid. HDAC8 showed a preference for the carboxylamine peptide by ~2.2-fold. While this difference is relatively insignificant, the next set of peptides were purchased with a carbamide at the C-terminus.

## **FlexPepBind protocol**

We have adapted the Rosetta FlexPepBind protocol[8, 9] to predict the substrate specificity of histone deacetylase 8 (HDAC8). The core strategy of the protocol is to model a peptide-HDAC8 complex structure for each of the investigated peptide sequences, and then to distinguish binders from non-binders based on these structural models. This involves calibration of different parameters such as the template structure used to represent the receptor (HDAC8), the peptide, the amount of sampling, and the adaption of the scoring function. Conformational sampling is biased towards relevant conformations by implementation of constraints that reproduce conserved structural features identified from known structures of the interaction, or from experiment.

The underlying assumption of this approach is that chemistry is the kinetic mechanism for steady state turnover is in rapid equilibration with substrate to form the ES complex followed by equilibration with substrate followed by rate-limiting chemistry (see chapter 5 for an explanation). This is a reasonable assumption given the large  $K_M$  and low  $k_{cat}$  values for peptides. As a result, utilizing binding data to model the catalytically activated complex ( $k_{cat}/K_M$ ) would inform HDAC8 substrate preference. Our goal was therefore to obtain maximal separation between peptides that are HDAC8 substrates (i.e., binders) and those that are not (i.e., cannot bind). We also sought to define an algorithm which produces a good correlation between predicted relative binding affinities and corresponding experimental values. We first optimized the protocol on a training set of peptides that have been tested for deacetylation catalyzed by HDAC8. Subsequently, we validated our protocol on a corresponding test set, and a number of further validation sets of peptides that were analyzed as we proceeded with the protocol development. Finally, the calibrated protocol was applied to identify new substrates. In the following, we describe the different steps of this strategy in detail.

## **Modeling the structures of peptide-protein complexes**

A major component of our protocol is the generation of a peptide-receptor complex structure for each of the investigated peptides. The first step involves the generation of a coarse starting



structure of the peptide-receptor complex. In the second step, this starting structure is then optimized, by either (a) extensive optimization with Rosetta FlexPepDock[10], or (b) simple minimization of the peptide-receptor interface. In the third step the optimized structure is scored, and peptides are ranked for their ability to bind the HDAC8 activated complex.

### **Step 1: Preparation of a coarse starting model of a peptide-HDAC8 complex**

For each of the peptides, a coarse model of the complex was generated, based on a template crystal structure of HDAC8 bound to a peptide substrate, and the sequence of the given peptide. PDB ID 2v5w[11] was used as a template: it includes the catalytically dead mutant Tyr306Phe HDAC8 in complex with a p53-derived acetylated peptide Ac-RHKacKac-methylcoumarin, and is expected to represent efficient peptide binding.

We tested 3 approaches to create the starting complex: (1) Initially, we created an extended peptide starting from the acetylated Lysine in the peptide of the solved crystal structure (all  $\Phi$  angles were set to -135.0 degrees, all  $\Psi$  angles to +135.0 degrees). Subsequently, we created a better starting structure, by (2) using the peptide backbone in the solved structures where available, (and complementing the missing part with an extended conformation), and finally (3) using extensive optimization of the peptide backbone conformation for an efficient HDAC8 peptide substrate (using Rosetta FlexPepDock[10], see below). The latter assumes that most peptides will adapt a similar conformation in the HDAC8 binding site, and that this conformation will be easiest to model for a high affinity substrate onto HDAC8.

We threaded different peptide sequences, to create starting structures for further optimization (below). Rosetta fixed-backbone design[12] was used to thread the peptide sequence onto the template peptide backbone conformation.

## **Step 2: Optimization of the peptide-HDAC8 complex structure**

In our previous studies on peptide substrate specificities using FlexPepBind[8, 9], we have used two main approaches for the optimization of a peptide-receptor structure from the coarse starting model, namely: (a) Extensive optimization using Rosetta FlexPepDock[10, 13] and (b) Short minimization of the starting structure. Minimization was only shown in previous studies of FlexPepBind to provide results of similar quality as the extensive FlexPepDock optimization protocol for calibration for FTase peptide substrate specificity[8], while for BCL specificity, full optimization was critical for good performance[9].

### **Step 2(a) Extensive optimization using Rosetta FlexPepDock**

This protocol has been developed previously in our lab to robustly refine coarse models of peptide–protein complexes into high resolution models. Optimization includes all internal degrees of freedom of the peptide (backbone and side chain torsion angles), as well as its rigid body orientation relative to the receptor (translation and rotation), and the side chains of the receptor. The receptor backbone is fixed. To allow significant perturbations within the binding pocket, while preventing the peptide and protein from separating during energy minimization, the simulation initially begins with strongly reduced (2 %)/increased (225 %) weights of the repulsive/attractive Van der Waals terms, which are gradually ramped back towards their original weights during the optimization. This non-deterministic protocol is applied repeatedly to generate a set of models (typically 200; 1000 models were generated for optimization of the starting structure, see above), and the best-scoring model is selected (according to a scoring function detailed below).

## Step 2(b) Short minimization of starting structure

We also investigated a shorter protocol that simply minimizes all internal degrees of freedom of the peptide, the receptor side chains, and the rigid body orientations (as implemented for FTase specificity prediction[8]).

## Parameters optimized for HDAC8 FlexPepBind

### Biasing conformations with constraints

Conformational sampling was biased towards relevant conformations by implementation of constraints that reproduce conserved structural features identified from known structures of the interaction or experimental data. HDAC8 catalyzes deacetylation of a wide range of different substrates, which all must bind to the catalytic active site. Maintaining the structural features that are important for binding is critical to structure-based substrate specificity prediction. Inspection of HDAC8 bound to a methylcoumarin conjugated peptide representing p53 (PDB ID 2v5w), identified potential critical interactions that we implemented as constraints. These constraints, and how they are implemented into the Rosetta scoring function, are listed in Table 3.1 and shown in Figure 3.1.

	Constraint type	Function type	Atoms involved	Constraint
1	Atom Pair <sup>a</sup>	BOUNDED <sup>b</sup>	D101 O <sup>δ2</sup> : Kac N	2.2 - 3.2Å
2	Atom Pair	BOUNDED	D101 O <sup>δ1</sup> : Coumarin N	2.5 - 3.5Å
3	Dihedral <sup>c</sup>	CIRCULAR HARMONIC <sup>d</sup>	Kac N : Kac C <sup>α</sup> : Kac C : Coumarine N	5.46 radians
4	Atom Pair	HARMONIC <sup>e</sup>	D267 O <sup>δ2</sup> / D267 O <sup>δ2</sup> : Kac O <sub>n</sub>	2.8Å

5	Atom Pair	HARMONIC	H180 N <sup>δ1</sup> : Kac O <sup>n</sup>	3.3Å
6	Atom Pair	HARMONIC	G151 O : Kac N <sup>κ</sup>	3.0Å
7	Atom Pair	HARMONIC	F152 C <sup>ε1</sup> / F152 C <sup>ε2</sup> : Kac C <sup>δ</sup>	3.8Å
8	Atom Pair	HARMONIC	F208 C <sup>δ1</sup> / F208 C <sup>δ2</sup> : Kac C <sup>γ</sup>	3.7Å

**Table 3.1 Different constraints derived from the solved structure ( 2v5w )**

a Distance constraint between two atoms

b Bounded function:

$$f(x) = \begin{cases} \left(\frac{x-ub}{ub-lb}\right)^2 & \text{for } x < lb \\ 0 & \text{for } lb \leq x \leq ub \\ \left(\frac{x-lb}{ub-lb}\right)^2 & \text{for } ub < x \leq ub + rswitch \cdot sd \\ \frac{1}{2}(x - (ub + rswitch \cdot sd)) + \left(\frac{x-ub}{ub-lb}\right)^2 & \text{for } x > ub + rswitch \cdot sd \end{cases}, \text{ where } lb, ub$$

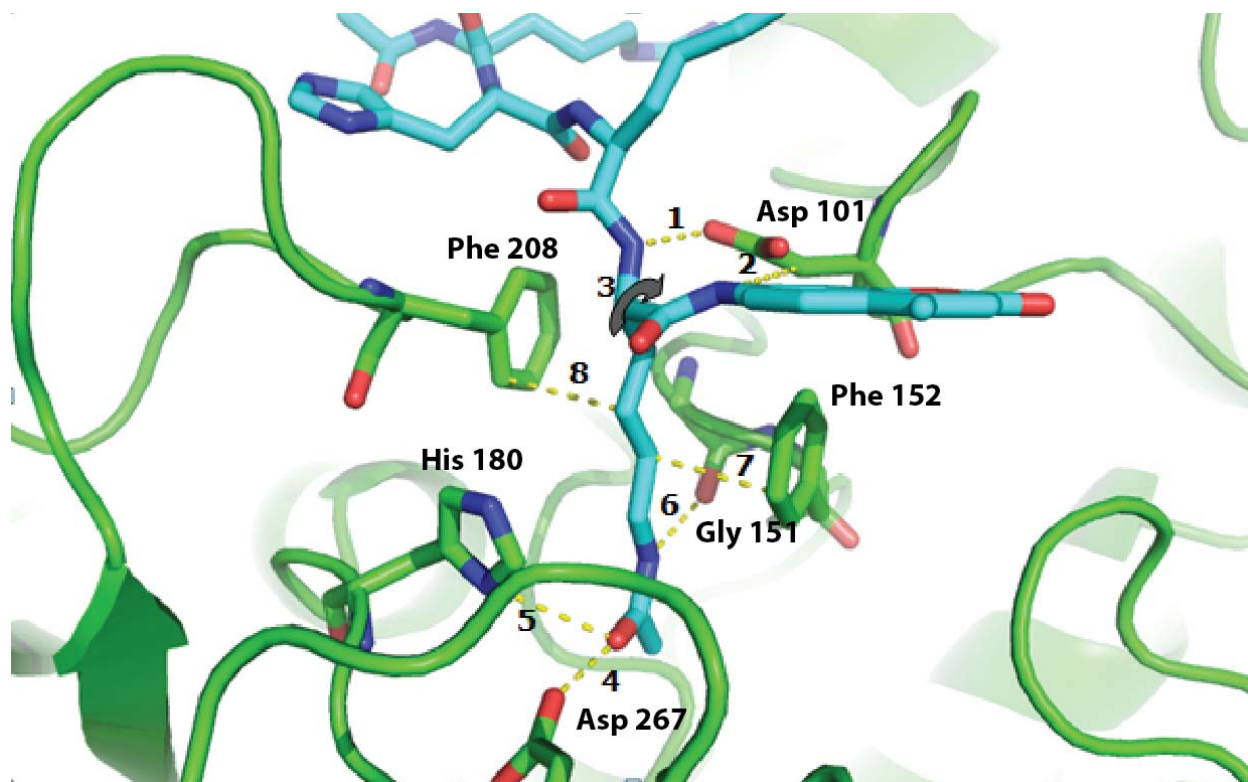
represents the boundaries within which zero scoring penalty is applied. rswitch is set to 0.5

c Dihedral constraint between four atoms

d Circular harmonic function:  $f(x) = \left(\frac{\text{Nearest Angle Radians}(x, x0) - x0}{sd}\right)^2$ , where  $x0$  and  $sd$  represent the radian in the starting structure and a standard deviation (set to 0.2), respectively.

e Harmonic function:  $f(x) = \left(\frac{x - x0}{sd}\right)^2$

The first three constraints tether the peptide backbone of the acetylated lysine to the receptor via two hydrogen bonds to the side chain of receptor residue Asp101. This enforces a cis dihedral angle conformation between the two adjacent backbone nitrogen atoms. The importance of



**Figure 3.1 Key interactions within the HDAC8-substrate interface**

Receptor in green, peptide in cyan, peptide residues and interacting residues in the receptor are shown as sticks.

Asp101 is evident from the fact that the Asp101Ala mutation leads to complete loss of deacetylation activity with both peptides and purified histones substrates[11]. The next two constraints tether the acetyl group of the lysine to the zinc ion (using zinc-binding residues His180, Asp267; the zinc atom itself is not included in the simulations). The remaining constraints, interaction of the Gly151 oxygen atom with the lysine nitrogen atom, and the stacking interactions of the side-chain rings of Phe152 and Phe208 residues with the acetylated lysine side-chain, tightly lock the acetylated lysine side-chain within the binding pocket.

### **Definition of axes for rigid body movements**

The rigid body orientation between the HDAC8 receptor and the peptide substrate is defined by two anchors that connect the receptor to the peptide. In the FlexPepDock framework, these

anchors are by default defined as the peptide  $\alpha$  carbon atom nearest to the peptide center of mass, and the receptor  $\alpha$  carbon atom nearest to the peptide anchor. In order to optimize rigid body sampling to relevant conformations, we explicitly defined the  $\alpha$  carbon atom of the acetyl lysine as the peptide anchor.

### **Addition of an electrostatic term to default scoring function score12**

Initially, we used the Rosetta default scoring function in all simulations (score12). Subsequently we added a weak, short range Coulombic electrostatic energy term to the scoring function (fa\_elec), which consists of a simple, linearly increasing and distance dependent dielectric term to model solvent screening effects, with all interactions truncated at 5.5 Å. This preserves the short-ranged nature of the all-atom potential. In addition, the penalty for burial of a carboxyl group oxygen atom was increased to prevent formation of too many buried polar interactions (the  $\Delta G_{\text{free}}$  parameter of the Lazaridis-Karplus solvation potential was modified from -10.0 to -13.5 for this atom type). In a previous study, we found these parameters improve our prediction of BCL binding specificity[9]. Moreover, an independent study on DNA binding specificity demonstrated that the incorporation of this explicit electrostatics term, together with the modification of one of the parameters in the Lazaridis-Karplus solvation potential, in addition to Rosetta's orientation-dependent hydrogen bonding potential[14], helped to prevent unfavorable short-range electrostatic interactions, modulated the interaction strength of charged and polar hydrogen bonds, and generally improved the performance of their DNA-protein interaction specificity predictions[15].

### **Step 3: Ranking of the peptide substrates based on the optimized peptide-receptor complex structures**

In the last step of the FlexPepBind protocol, different peptide-HDAC8 activated complexes are ranked to model the relative binding ability of different peptides to HDAC8 (and thus, their ability to undergo deacetylation catalyzed by HDAC8). In order to focus on the peptide-HDAC8 interface, different subsets of the total score were investigated:

(1) Peptide score - includes only energy terms that involve the peptide, i.e. the internal energy of the peptide, and interactions across the interface. A similar term, Peptide score (noref) was used in the FlexPepBind FTase specificity protocol[8]. This term does not contain the reference energy term, (Eref), originally introduced to allow comparison of structures of different sequence, and to bias for native protein sequences during fixed backbone design.

(2) Interface score - includes the sum of interactions across the interface.

(3) Reweighted score - sum of peptide score, interface score and total score. This score gives more weight to the contribution of peptide to the overall energy.

### **Dataset 1: Training and test set: 361 GXKacZGC peptides**

The Fierke group has tested the ability of HDAC8 to deacetylate 361 6-mer peptides with the sequence Ac-GXKacZGC-NH<sub>2</sub> (where X,Z can be any amino acid, except cysteine), under two different conditions: for zinc and iron bound HDAC8 (Joseph and Fierke unpublished results). For each of these peptides, the level of deacetylation by HDAC8 was measured by reacting the substrate for 30 minutes at 30°C, for both Zn- and Fe-bound HDAC8 and measuring the fraction product with a SAMDI mass spectrometric assay in collaboration with the Mrksich group. In this study, we focus only on Zn bound (0.5 μM Zn(II)-HDAC8) as a first step in deciphering the entire co-factor dependent specificity profile of the enzyme. We divided this dataset of 361 peptides into a training and a test set, by sorting the peptides according to their experimental activity with Zn(II)- HDAC8, and assigning all even/odd-numbered rows to be the test/training set, respectively. This division assured an even distribution of peptides with respect to their activity levels (avoiding a situation where one set holds a large number of high/low activity substrates).

### Initial calibration set: 5 substrates & 5 non-substrates

As a first step, we evaluated how well a protocol can distinguish between 5 active substrates and 5 peptides that are not HDAC8 substrates (Table 3.2).

Substrates	Non-substrates
GYKacFGC (93%)	GDKacRGC
GYKacWGC (80%)	GYKacEGC
GLKacFGC (66%)	GNKacHGC
GFKacFGC (64%)	GTKacKGC
GIKacFGC (62%)	GMKacPGC

**Table 3.2 Initial calibration set**

### Dataset 2: Validation set: 361 GRKacXZC peptides

Mrksich *et al.* evaluated the reactivity of HDAC8 with a library of peptides (Ac-GRKacXZC-NH<sub>2</sub>) peptides with many possible amino acid combinations of the X and Z positions[2]. The arrays were prepared by immobilizing the peptides to a self-assembled monolayer of alkanethiolates on gold and then analyzed by a mass spectrometry technique termed SAMDI (self-assembled monolayers for matrix assisted laser desorption/ionization time-of-flight mass spectrometry).

### Statistical tests

In this study, we used two measurements to evaluate the performance of HDAC8 FlexPepBind substrate prediction: (1) Binary distinction between substrates and non-substrates, and (2) Linear correlation between predicted relative binding affinity and experimentally determined  $k_{cat}/K_M$  values for each substrate. An optimal protocol should be able to reliably identify (at least the strongest) peptide substrates, and preferably to rank peptide substrates according to their relative reactivity with HDAC8.



## **Binary distinction between substrates and non-substrates**

For binary distinction, we determined a deacetylation level, beyond/below which a peptide is defined as a substrate/non-substrate, respectively (below). The non-parametric Kolmogorov-Smirnov test was applied to determine whether substrates and non-substrates are derived from two distinct populations (e.g., they can be distinguished).

## **Prediction of relative substrate activity**

In order to assess the ability of FlexPepBind to rank substrate peptides according to their reactivity with HDAC8, we calculated correlations and the associated statistical significance, using the non-parametric spearman correlation test.

False Discovery Rate (FDR) corrections were applied to evaluate the statistical significance taking into account the number of hypotheses tested for each data set.

## **Results**

### **First round of calibration**

After development of an initial algorithm based on the library data, we decided to perform a quick calibration on a small set of peptides (5 peptide substrates proposed to have high HDAC8 activity and 5 peptides proposed to have low HDAC8 activity) (Table 3.2). A satisfactory distinction between substrates and non-substrates could be obtained in the computational model using: (1) Optimization of structure with minimization only, rather than the more time consuming FlexPepDock optimization; (2) Default scoring using score12; (3) The acetylated lysine substrate constrained to its present location using a set of distance and dihedral constraints (Table 3.3; standard deviation of constraints: 0.2); (4) The structural template HDAC8-peptide: 2v5w with an extended peptide conformation (see Methods); and (5) Defined peptide anchor

residue of  $\alpha$  carbon of the acetylated lysine (residue 366 in HDAC8 structure). These initial parameters were chosen based on experience in previous studies[8, 9], and/or from inspection of the solved HDAC8-peptide complex structure (PDBID 2v5w). Table 3.3 shows that substrates could well be distinguished from non-substrates, using any of the scores for ranking.

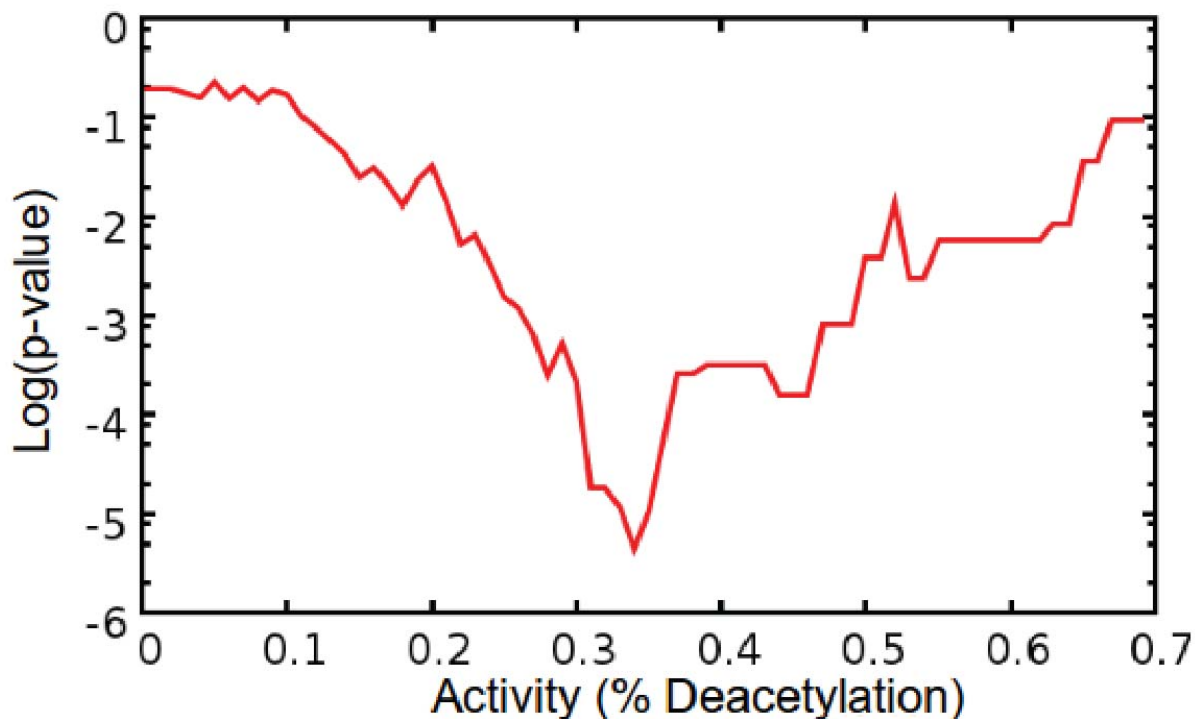
Scoring scheme	Interface score	Peptide score	Peptide score no ref	Reweighted score
KS p-values	0.036	0.036	0.036	0.036

**Table 3.3 Distinction between substrates and non-substrates in the initial calibration run**

For the next phase, we ran our peptide modeling protocol on the whole training set, mainly using the set of parameters that exhibited superior performance in this short calibration phase.

### **Analysis of performance in Training and Test sets**

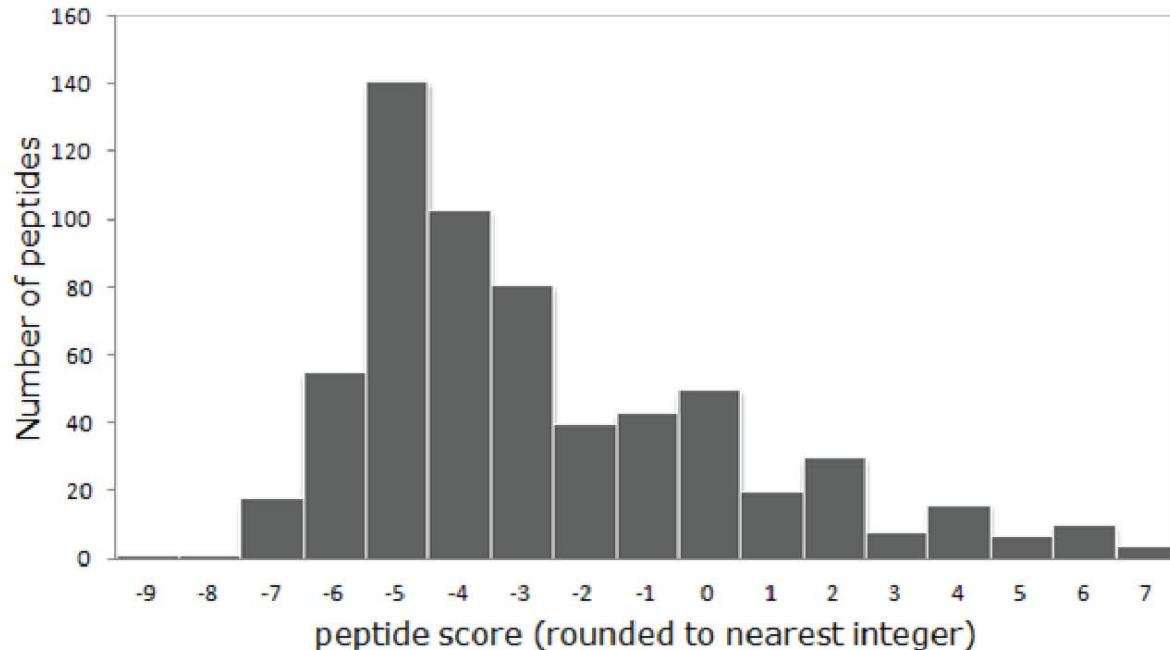
The original dataset measured by Caleb Joseph contains sequences of acetyl lysine peptides that are ranked by their activity level as substrates, we needed to develop a protocol for binary classification into substrates and non-substrates with a cutoff value that can be used as a lower limit to define peptides as substrates. We derived that cut off by applying a 2 sample Kolmogorov-Smirnov (KS) test on all possible activity levels ([0,1], in resolution of 0.01). The activity level that was chosen as a cutoff is the one that obtained the lowest p-value in the KS test, thus, the one that could best differentiate between the 2 distributions of scores for substrates vs. non substrates (Figure 3.2). Our best distinction between substrates and non-substrates is obtained for a cutoff of 34 % deacetylation under defined assay conditions. This value did not change significantly upon variation of parameters (data not shown).



**Figure 3.2 Log(p-value) of KS test vs. activity level plot shows that the best cutoff for binary classification is around 0.34**

### Searching for novel, non-histone substrates

The Phosphosite database[16] (<http://www.phosphosite.org/>) contains a compilation of all experimentally observed acetylation sites in proteins. We extracted from this database 3184 acetylated sites in 1603 proteins. These present a pool of potential targets for catalysis of deacetylation by HDAC8. In order to use our algorithm to predict their ability to be deacetylated by HDAC8, we shortened the sequences around the acetylated lysine to the same size of the sequences in our experimental dataset - `XXKacXXX`, and used these as input. Figure 3.3 shows the distribution of peptide scores for the 3184 acetylation sites. From the previous training and test runs, we defined a threshold value of peptide score -5 units, above which our confidence in a true prediction is very high. While most sequences in the Phosphosite database are probably not highly efficient HDAC8 substrates, we predict that some peptides with scores lower than -5 units will nevertheless be substrates. This finding could suggest that there are quite a number of potential substrates of HDAC8 or other deacetylases that are yet to be discovered.



**Figure 3.3 Distribution of peptide scores in the 3184 acetylated lysine sequences**

The rightmost bar concentrates all the peptides that have a minimization score below 7 (a high score that suggests that our approach was not able to model successfully these peptides into the binding site).

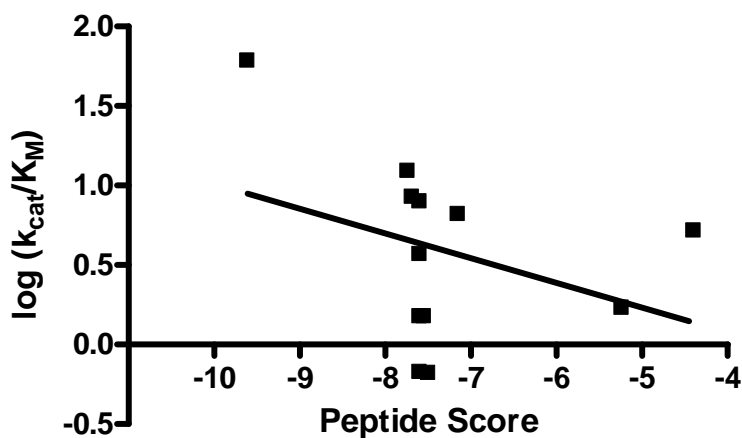
### Experimental validation of de-novo predictions

To test our prediction algorithm, we selected a set of 26 acetylated peptide sequences for experimental validation. This set includes:

(1) A positive set (“Good” peptide) - the 10 top-ranking peptides in phosphosite (peptide score  $\leq -6$ ). In addition to this positive set, we selected acetylated peptides that occur in the same proteins: (2) A negative set (“bad” peptides) - 10 peptides predicted not to undergo deacetylation by HDAC8 (high peptide scores 0 to 7), and (3) Additional potential substrates that ranked in the middle (-6 to 0) (“Mediocre” peptides). 20 out of 26 peptides were successfully characterized by experiments. This set of 20 peptides is named Pep20 set throughout the rest of the paper.

In order to assess the accuracy of the algorithm, the predicted peptides were purchased and assayed. These peptides were dissolved in 50% acetonitrile, and 20 of the 26 peptides were

soluble. The zinc content of one of the peptides (Ac-VSKacGPF-COO<sup>-</sup>) was determined, using ICP-MS, to be one-one hundredth of the peptide concentration. All soluble peptides were assayed using a coupled assay developed to measure the acetate release from a deacetylation reaction (Table 3.4) and varied three orders of magnitude in their initial velocities. Of these peptides, the peptide representing the Lys350 of the histone acetyl transferase KAT6A (Ac-VSKacGPF-COO<sup>-</sup>), was deacetylated four-fold faster than the next fastest peptide ( $61 \text{ M}^{-1}\text{s}^{-1}$  versus  $12 \text{ M}^{-1}\text{s}^{-1}$ ). Two peptides representing other acetylation sites on KAT6A (Lys604 and Lys 815) were also assayed. These peptides were deacetylated by HDAC8 slower than the Lys350 peptide ( $7.5 \text{ M}^{-1}\text{s}^{-1}$  and  $> 0.6 \text{ M}^{-1}\text{s}^{-1}$ , respectively). Finally, to determine the accuracy of the Score12, the  $\log(k_{\text{cat}}/K_{\text{M peptide}})$  was graphed versus the Score12 score (Figure 3.4). This comparison yielded a direct linear correlation with an  $R^2$  value of 0.15.



**Figure 3.4 Experimental validation of the Pep20 set**

Experimental values (X-axis) plotted against the predicted score (Peptide score, Y-axis, all positive scoring peptides have been assigned zero score for better representation) indicate that while strong substrates are favorably scored, this prediction scheme also identifies many false positive hits (points in the lower left of the figure).

Peptide Sequence	Uniprot Accession	Protein	Kac position	Peptide score	$\log(k_{cat}/K_M)$
<b>SGKacKGQ</b>	Q92922	SWI/SNF complex subunit SMARCC1	345	-7.5	0.66
<b>IIKacDGE</b>	O95831	Apoptosis-inducing factor 1	593	-7.55	1.5
<b>DIKacYPL</b>	P00387	NADH-cytochrome b5 reductase 3	42	-7.6	7.9
<b>PGKacGVK</b>	Q01094	Transcription factor E2F1	117	-7.6	3.7
<b>DDKacYTL</b>	Q15005	Signal peptidase complex subunit 2	169	-7.6	0.67
<b>TQKacQEQ</b>	P12270	Nucleoprotein TPR	755	-7.6	1.5
<b>VSKacGTL</b>	P16401	Histone H1.5	93	-7.69	8.5

<b>SFKacSDQ</b>	Q9UPN9	E3 ubiquitin-protein ligase TRIM33	769	-7.74	12
<b>ESKacFQQ</b>	P06400	Retinoblastoma-associated protein	896	-8.79	
<b>VSKacGPF</b>	Q92794	Histone acetyltransferase KAT6A	350	-9.61	61
<b>Predicted non-substrates in same proteins</b>					
<b>FSKacVRT</b>	Q92794	KAT6A	355	1792	
<b>VGKacSVS</b>	Q92794	KAT6A	815	1751	<0.6
<b>DHKacTLY</b>	Q92794	KAT6A	604	174	7.5
<b>MIKacHLE</b>	P06400	Retinoblastoma-associated protein	548	186	
<b>FTKacDHL</b>	Q9UPN9	TRIM33	252	965	3.2
<b>KGKacTAQ</b>	Q9UPN9	TRIM33	953	2.61	3.1
<b>GVKacKVA</b>	P16401	Histone H1.5	168	624	1
<b>KPKacAAK</b>	P16401	Histone H1.5	209	145	4.1
<b>SFKacLNK</b>	P16401	Histone H1.5	109	16.17	0.061
<b>IVKacEVE</b>	P12270	Nucleoprotein TPR	428	1430	

<b>NQKacLTA</b>	P12270	Nucleoprotein TPR	748	975	<0.6
<b>Additional predicted substrates</b>					
<b>VSKacAVE</b>	P06733	Alpha-enolase	64	-5.07	
<b>MGKacGVS</b>	P06733	Alpha-enolase	60	-5.16	
<b>SGKacYDL</b>	P06733	Alpha-enolase	256	-7.15	6.6
<b>DSKacNAK</b>	Q92793	CREB Binding Protein CBP	1583	-4.4	5.2
<b>KKKacNNK</b>	Q92793	CBP	1588	-5.24	1.7

**Table 3.4 Experimental measurement of potential substrates and a corresponding predicted negative set**

These experimental results show that while our initial protocol (namely minimization only, using the default scoring function (score 12), and starting from an extended peptide) performs well on training and test sets the ability to predict HDAC8 substrates as indicated by the Pep20 results: While strong substrates are recognized very well, the overall correlation between the predicted score and experimental activity is low, largely due to false positives (Figure 3.4 and Table 3.4).

### **Second round of calibration**

The reduced performance on the Pep20 set compared to training and test sets indicates that the training set was too restricted in sequence variability and that a more general protocol might be developed by including more diverse substrates. This motivated us to perform a second round of calibration, starting from Pep20. Our aims were to improve performance on the Pep20 set, while



maintaining similar performance on the previous training and test sets. Furthermore, we were interested in maintaining good performance with a more extended, fully optimization protocol, to ensure robustness, and to enable the modeling of very different peptide substrates in the future.

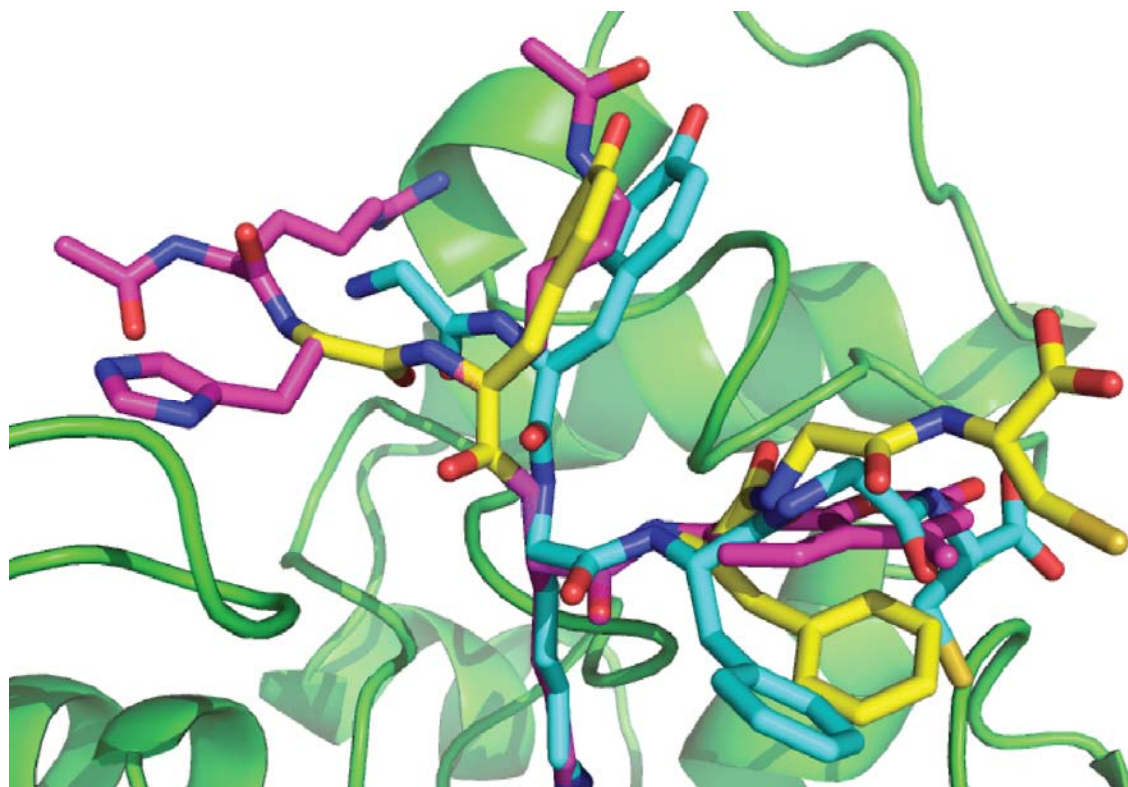
For the second round of calibration, we investigated three main parameters of our model, each relating to a different step of the protocol. In step 1, preparation of starting structure, we attempted to create a better starting structure, assuming that this would result in better predictions. In step 2, optimization of peptide-HDAC8 complex structure, we investigated inclusion of an electrostatic term into the default scoring function, motivated by improved performance observed in previous studies on prediction of BCL peptide binding specificity[9] and DNA binding specificity[15]. In step 3, in which different peptide substrates are ranked, we further evaluated different scoring terms, in particular Interface score vs. peptide score, which were both shown to perform similarly in the first optimization round, but produced different top-ranking substrate lists.

### **Improved starting structure and score function**

While the template in the initial protocol contained an extended peptide conformation anchored on the acetylated lysine substrate residue, we investigated how to include more information about peptide backbone conformations preferred by HDAC8. Since the only peptide substrate (Ac-RHKacKac-methylcoumarin) solved with HDAC8 (PDB ID 2v5w; see Methods) does not cover the full peptide backbone of the hexamer peptides used to calibrate the protocol (XXKacZZZ), atoms beyond the nitrogen atom of the methylcoumarin need to be modeled.

We investigated two possible approaches: (1) partial threading of the peptide residues combined with an extended conformation (XXKacZZZ, X,Kac=threaded, Z extended; see Methods), and (2) an optimized template. The latter was obtained starting from a partially threaded peptide (the first option above), onto which the sequence of an active substrate, Ac-GYKacFGC-NH<sub>2</sub> was threaded. This structure was then optimized using the full optimization protocol (see Methods). The underlying assumption is that a backbone conformation optimized for a sequence known to

be a good substrate is a better starting structure for modeling peptide sequences into the HDAC8 substrate binding site. These two starting structures are shown in Figure 3.5.



**Figure 3.5 Improved starting structures for optimization**

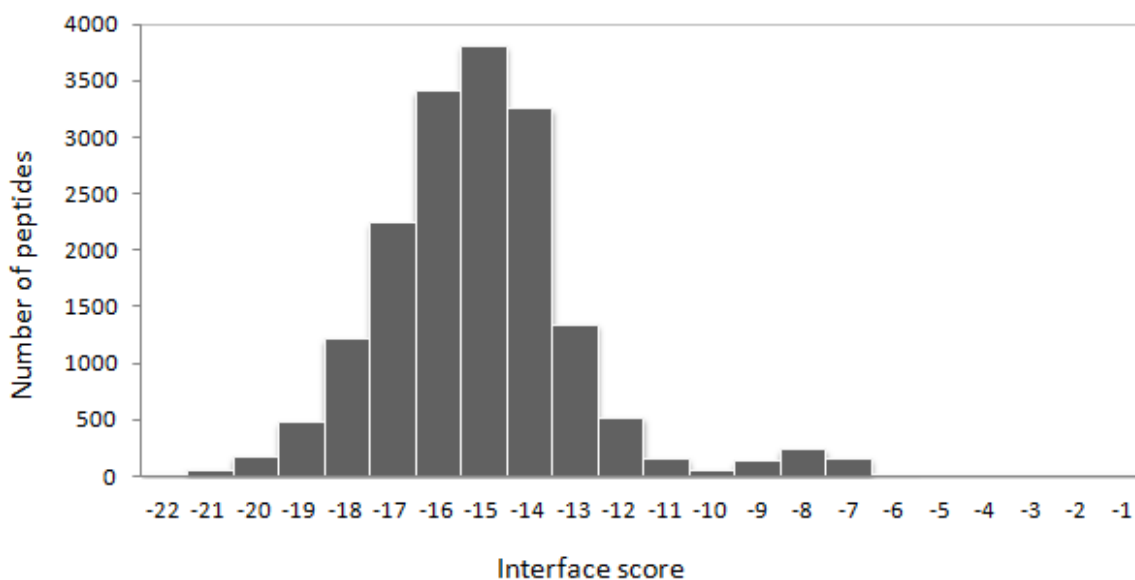
In green: HDAC8 crystal structure (2v5w); in yellow: partially threaded template; cyan: optimized template). Note that while the acetyl lysine is not changed, the position of the N- and C-terminal peptide tails vary.

We also increased the number of design runs ( $\text{ndrun} = 100$ ), to obtain better starting structures (see Methods). The optimized template significantly improved the prediction of substrate efficiency in the Pep20 set. We therefore chose this template as a starting structure for the further optimization steps below.

To further improve our protocol, we investigated modifications of the scoring function. We improved prediction of HDAC8 substrates by inclusion of electrostatic energy term.

## Second round of Phosphosite screen

With this improved protocol in hand, we proceeded to a second screen of the Phosphosite database. For this, we used the improved protocol after trimming sequences around the acetylated lysine to hexamers, as in our experimental datasets. The goal was to screen for putative novel HDAC8 targets by top-ranking peptides. Peptides were again chosen as “good” (interface score  $< -17$ ), “bad” (interface score  $> -12$ ) and “mediocre” ( $-12 >$  interface score  $> -17$ ) substrates (Figure 3.6). We therefore compiled Pep19, a set of 19 peptides (Table 3.5) that could help us to validate the new protocol along with choosing the best scoring term for ranking and discrimination between substrates and non-substrates.



**Figure 3.6 Distribution of interface score in the acetylated peptides**

## Assessing the second set of peptides

A second set of 19 peptides (17 of which were soluble) were purchased and rendered metal free using Chelex 100 resin. Zn(II)-HDAC8 and two peptide concentrations ( $\geq 100 \mu\text{M}$ ) were incubated and measured using the coupled assay system (Chapter 2). The initial slopes were fit to a line yielding  $k_{\text{cat}}/K_{\text{M}}$  values. The catalytic efficiencies of these peptides range over 1000-fold ( $< 2 \text{ M}^{-1}\text{s}^{-1}$  to  $4,400 \text{ M}^{-1}\text{s}^{-1}$ ). Peptides corresponding to a DNA mismatch repair protein Msh6, E3 sumo ligase CBX4, and Elongation factor 1-alpha 1 are the fastest measured peptides (to date) for HDAC8 catalyzed deacetylation with rates of  $880 \text{ M}^{-1}\text{s}^{-1}$ ,  $1,200 \text{ M}^{-1}\text{s}^{-1}$ , and  $4,400 \text{ M}^{-1}\text{s}^{-1}$  respectively. In this case, there is a rough linear correlation between  $\log(k_{\text{cat}}/K_{\text{M}})$  and the Interface score (Figure 3.7). This graph also resulted in a linear correlation with an  $R^2$  value = 0.4 which is better than double  $R^2$  the value for the pep20 set.

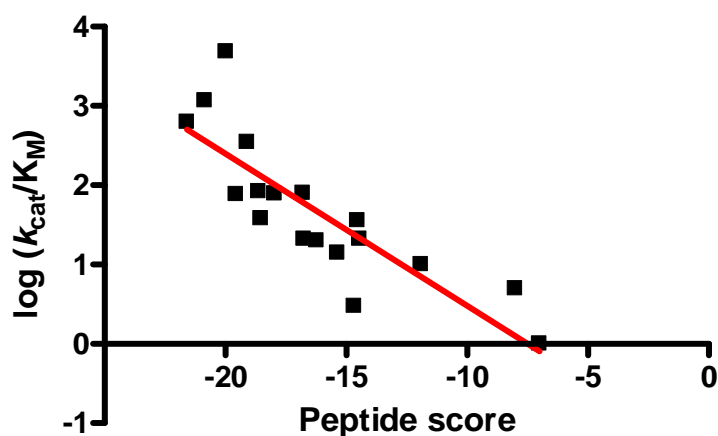
For better insight into the recognition of efficient peptide substrates by HDAC8, Michaelis-Menten kinetics were determined for the three fastest peptides.  $1 \mu\text{M}$  Zn(II)-HDAC8 and 0 - 400  $\mu\text{M}$  peptides were incubated under standard assay conditions (Table 3.5). For these peptides, low  $k_{\text{cat}}$  values ( $< 0.05 \text{ s}^{-1}$ ) consistent with non-fluorophore conjugated peptides were measured (Appendix A).  $K_{\text{M}}$  ( $< 50 \mu\text{M}$ ) were measured and appear to be the deterring factor in the high  $k_{\text{cat}}/K_{\text{M}}$  values. The peptide concentration dependence (0-400  $\mu\text{M}$ ) for the initial rate for deacetylation catalyzed by HDAC8 was determined (Table 3.5). The catalytic efficiency ( $k_{\text{cat}}/K_{\text{M}}$ ) is mainly due to a decrease in  $K_{\text{M}}$ , likely reflecting enhanced binding interactions. This is consistent with the algorithm which mainly measures enhanced HDAC8-peptide affinity.

Peptide	Uniprot ID	Protein	Kac position	Interface score	peptide score	$k_{\text{cat}}/K_{\text{M}}$
SLKacEFY	P30838	Aldehyde dehydrogenase, dimeric, NADP-preferring	K269	-17.9	-7.8	78

<b>LSKacFLR</b>	P27708	CAD protein	K747	-21.6	-7.3	880
<b>ISKacYDR</b>	P52701	DNA mismatch repair protein Msh6	K504	-28.6	-8.2	85
<b>RKKacGEP</b>	P04637	Cellular tumor antigen p53	K292	-11.9	-4.9	13
<b>KIKacRLR</b>	P18124	60S ribosomal protein L7	K31	-18.5	-5	7.7
<b>LGKacFRR</b>	Q6PKG0	La-related protein 1	K1017	-19.1	-6.1	50
<b>RLKacYSQ</b>	Q14683	SMC1A	K713	-16.8	-7.2	83
<b>SGKacYFA</b>	Q01433	AMP deaminase 2	K517	-16.2	-9.2	20
<b>TWKacANF</b>	P14316	Interferon regulatory factor 2	K78	-19.6	-7.9	77
<b>TFKacGVD</b>	Q9UK76	Hematological and neurological expressed 1 protein	K8	-15.4	-8.2	14
<b>GGKacAFG</b>	P55008	Allograft inflammatory factor 1	K11	-14.5	-7.4	21

<b>FVKacAFA</b>	P51649	Succinate-semialdehyde dehydrogenase, mitochondrial	K358	-14.7	-8.9	2.8
<b>GIKacPFL</b>	Q9NQW6	Actin-binding protein anillin	K371	-7	-7.1	<2
<b>FGKacFSW</b>	Q5VUA4	Zinc finger protein 318	K1275	-19.9	-9.5	4400
<b>SFKacYAW</b>	P68104	Elongation factor 1-alpha 1	K55	-20.8	-10.2	1200
<b>SGKacYYY</b>	O00257	E3 SUMO-protein ligase CBX4	K149	-16.7	-9.7	21
<b>SEKacPEK</b>	P78527	DNA-dependent protein kinase catalytic subunit	K1970	-8	-2.8	21
<b>KGKacDAE</b>	Q9Y618	Nuclear receptor corepressor 2	K878	-14.5	-5	2.6

**Table 3.5 The second experimental validation: Pep19 set**



**Figure 3.7 Performance of new protocol on the second validation set, Pep19**

Data was fit to a linear equation yielding an  $R^2$  value of 0.7.

## Discussion

Determining HDAC substrate preference is integral to determining the role of HDAC8 in cellular homeostasis and HDAC related diseases. As a result, identification of *in vivo* HDAC substrates is currently one of the important challenges of the HDAC field. One of the most successful high throughput methods utilizes systematic changes to peptide substrates and relates those changes to full length protein substrates[1-4]. In this method, peptide arrays made of a 3- to 7-mer peptides are altered systematically in one or two amino acid positions. These studies have proposed HDAC preferences for various amino acids in specific peptide positions. For instance, these studies have indicated that HDAC8 prefers an aromatic amino acid in the +1 position in the context of a specific sequence. Similarly, these studies do not describe reactivity with peptides in other contexts indicating “context dependence” recognition: where multiple amino acid interactions contribute to specificity. While these experiments have been somewhat successful in identifying HDAC8 preference for peptides, they have thus far failed to identify *in vivo* full length protein substrates for HDAC8. We therefore initiated this study to determine whether HDAC peptide preference could be improved by utilizing a computational approach to aid in

evaluating specificity. If a computational model is able to identify better peptide substrates, this would provide evidence for multiple interactions contributing to selectivity.

In this chapter, we used an initial training set followed by two rounds of optimization to determine a general substrate binding algorithm to predict “good” peptide substrates for HDAC8. Initial experiments were based on a training set of peptides for which kinetics were experimentally determined (Joseph, Mrksich, and Fierke unpublished)[1]. These experiments utilized a mass spectrometry based assay to determine the deacetylation rates for a set of peptides with systematic amino acid changes in the -1 and +1 positions, and in the +1, +2 positions. This study showed that HDAC8 displays a preference for aromatic amino acids (specifically phenylalanine) in the +1 position.

The set of substrates identified using the first algorithm had low values of  $k_{cat}/K_M$  for deacetylation catalyzed by HDAC8 ( $\leq 61 \text{ M}^{-1}\text{s}^{-1}$ ). The low catalytic activity, especially when compared to the second set of peptides, likely suggest that none of the corresponding proteins are *in vivo* substrates for HDAC8. The activity of this peptide may also be decreased by the inclusion of a  $-\text{COO}^-$  at the C-terminus which likely lowered activity  $\sim 2$ -fold (data not shown) and by inhibition by Zn(II) contamination of the peptides that could decrease activity by 70%[7].

After multiple rounds of optimization of the computational method, a roughly linear correlation between the algorithm score and the experimental  $\log(k_{cat}/K_M)$  value was observed providing confidence that this method evaluates HDAC8-peptide affinity. In fact the second validation set yielded peptides with the fastest known catalytic efficiencies ( $4400 \text{ M}^{-1}\text{s}^{-1}$  and  $1200 \text{ M}^{-1}\text{s}^{-1}$ ) for Zn(II)-HDAC8 peptide substrates corresponding to the elongation factor 1-alpha 1 and the E3 SUMO-protein ligase CBX4 respectively. These sequences contain aromatic amino acids in both the +1 and +3 positions (Ac-FGKacFSW-NH<sub>2</sub> and Ac-SFKacYAW-NH<sub>2</sub>). Aromatic motifs in the +1 position of peptides has been previously observed to increase HDAC8 catalytic efficiency towards peptide substrates[1, 2]. Pi-pi interactions between Tyr100 within HDAC8 and a methylcoumarin moiety at the +1 position have been observed in high resolution HDCA8-peptide structures and are theorized to increase the HDAC8 catalytic efficiency. It is not yet clear what interactions the aromatic amino acid in the +3 position are making with the HDAC8 peptide binding grove. Interestingly, another peptide containing three aromatics (Ac-



SGKacYYY-NH<sub>2</sub>) was deacetylated with a catalytic efficiency of only 20 M<sup>-1</sup>s<sup>-1</sup>, indicating that multiple aromatic amino acid side chains are not sufficient for activity. Thereby suggesting the formation of specific peptide-protein interactions.

The prediction of faster substrates by the second algorithm suggests that more than 2 amino acids are required to define HDAC8 catalysis of peptide substrates. This is due to the fact that the initial algorithm had difficulty identifying fast substrates in the first peptide set, as it only contained data from two amino acid residues. Once consideration of all 6 amino acids was taken into account, the algorithm was able to better predict faster substrates.

Previously, this algorithm was used to determine the substrate preference of another promiscuous post translational enzyme, farnesyltransferase. In the case of farnesyltransferase, this algorithm was less effective as it could only yield yes or no answers[8] as opposed to the gradient of rates that it predicts in this chapter. One possible explanation for this is that  $k_{cat}/K_M$  includes multiple steps for FTase activity and does not correlate with the peptide affinity[17]. HDAC8, which utilizes a rapid equilibrium substrate binding model, reflects both chemistry and binding affinity in the value of  $k_{cat}/K_M$ . As the algorithm mainly reflects binding affinity, this is a better correlation with HDAC8 substrate specificity.

We are hopeful that these substrates will improve the *in vivo* identification of HDAC proteins substrates. We are currently in the process of working with the Holson group to determine whether the acetylation state of proteins corresponding to the efficient peptide substrates increase after addition of an HDAC8 specific inhibitor (for method see Chapter 4), which would suggest that deacetylation *in vivo* is catalyzed by HDAC8.

In this chapter, we have determined how short range interactions affect HDAC8 substrate recognition, but how and whether long range interactions affect substrate specificity needs to be investigated. As the HDAC8 putative substrate recognition motif extends tens of angstroms beyond the binding site of short peptide substrates, it appears likely that interactions within distal regions of the peptide binding region may also foster HDAC8 substrate recognition. The work in the chapter provides a framework by which we can use to begin to ask this question. In the future, experiments utilizing the combinations of “good,” “bad,” and “mediocre” peptides and

singly acetylated full length substrates (Chapter 5) can be performed to understand how the contributions of the short and long range interactions affect substrate recognition.

## **Bibliography**

- [1] Z.A. Gurard-Levin, K.A. Kilian, J. Kim, K. Bahr, M. Mrksich, Peptide arrays identify isoform-selective substrates for profiling endogenous lysine deacetylase activity, *ACS Chem Biol*, 5 (2010) 863-873.
- [2] Z.A. Gurard-Levin, J. Kim, M. Mrksich, Combining mass spectrometry and peptide arrays to profile the specificities of histone deacetylases, *Chembiochem*, 10 (2009) 2159-2161.
- [3] Z.A. Gurard-Levin, M. Mrksich, The activity of HDAC8 depends on local and distal sequences of its peptide substrates, *Biochemistry*, 47 (2008) 6242-6250.
- [4] D. Riestler, C. Hildmann, S. Grunewald, T. Beckers, A. Schwienhorst, Factors affecting the substrate specificity of histone deacetylases, *Biochem Biophys Res Commun*, 357 (2007) 439-445.
- [5] B.E. Schultz, S. Misialek, J. Wu, J. Tang, M.T. Conn, R. Tahilramani, L. Wong, Kinetics and comparative reactivity of human class I and class IIb histone deacetylases, *Biochemistry*, 43 (2004) 11083-11091.
- [6] N.A. Wolfson, C.A. Pitcairn, C.A. Fierke, HDAC8 substrates: Histones and beyond, *Biopolymers*, 99 (2013) 112-126.
- [7] S.L. Gantt, S.G. Gattis, C.A. Fierke, Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion, *Biochemistry*, 45 (2006) 6170-6178.
- [8] N. London, C.L. Lamphear, J.L. Hougland, C.A. Fierke, O. Schueler-Furman, Identification of a novel class of farnesylation targets by structure-based modeling of binding specificity, *PLoS computational biology*, 7 (2011) e1002170.
- [9] N. London, S. Gulla, A.E. Keating, O. Schueler-Furman, In silico and in vitro elucidation of BH3 binding specificity toward Bcl-2, *Biochemistry*, 51 (2012) 5841-5850.
- [10] B. Raveh, N. London, O. Schueler-Furman, Sub-angstrom modeling of complexes between flexible peptides and globular proteins, *Proteins*, 78 (2010) 2029-2040.

- [11] A. Vannini, C. Volpari, P. Gallinari, P. Jones, M. Mattu, A. Carfi, R. De Francesco, C. Steinkuhler, S. Di Marco, Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex, *EMBO Rep*, 8 (2007) 879-884.
- [12] B. Kuhlman, G. Dantas, G.C. Ireton, G. Varani, B.L. Stoddard, D. Baker, Design of a novel globular protein fold with atomic-level accuracy, *Science*, 302 (2003) 1364-1368.
- [13] B. Raveh, N. London, L. Zimmerman, O. Schueler-Furman, Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors, *PLoS One*, 6 (2011) e18934.
- [14] T. Kortemme, A.V. Morozov, D. Baker, An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein-protein complexes, *J Mol Biol*, 326 (2003) 1239-1259.
- [15] S.B. Thyme, D. Baker, P. Bradley, Improved modeling of side-chain--base interactions and plasticity in protein--DNA interface design, *J Mol Biol*, 419 (2012) 255-274.
- [16] P.V. Hornbeck, I. Chabra, J.M. Kornhauser, E. Skrzypek, B. Zhang, PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation, *Proteomics*, 4 (2004) 1551-1561.
- [17] K.T. Lane, L.S. Beese, Thematic review series: lipid posttranslational modifications. Structural biology of protein farnesyltransferase and geranylgeranyltransferase type I, *Journal of lipid research*, 47 (2006) 681-699.

## Chapter 4

### Mining the Acetylome: An Unbiased Approach to Identify the Endogenous Substrates of Histone Deacetylase 8 (HDAC8)<sup>11,12,13</sup>

#### Introduction

Despite being extensively characterized structurally and biochemically, the functional role of histone deacetylase 8 (HDAC8) has remained largely obscure due in part to a lack of known cellular substrates. Herein, we describe an unbiased approach using chemical tools in conjunction with sophisticated proteomic methods to identify novel non-histone substrates of HDAC8. These proteins are involved in diverse biological processes such as mitosis, transcription, chromatin remodeling, and RNA splicing.

#### Results and discussion

Posttranslational acetylation of lysine residues is a highly conserved modification[1-3] which alters protein functions ranging from stability to activity and results in effects ranging from cytoskeletal reorganization[4] to changes in gene expression[5]. Histone deacetylases (HDACs) play a key role in maintaining the balance of acetylation states by catalyzing the removal of acetate groups from the  $\epsilon$ -amino position of acetyl lysine residues. As a result, these enzymes

---

<sup>11</sup> Reproduced in part from Olson, D.E. Udeshi, N. Wolfson, N.A. Pitcairn, C.A. Sullican, E.D. Jaffe, J. Scinkina, T. Natoli, T. Liu, X. Paulk, J. McCarren, P. Barker, D. Howe, E. Lazzaro, F. Zhang, Y. Gale, J. Subramanian, A. Fierke, C.A. Holson, E.B. 2014. Mining the Acetylome: An Unbiased Approach to Identify the Endogenous Substrates of “Histone” Deacetylase 8 (HDAC8). in preparation.

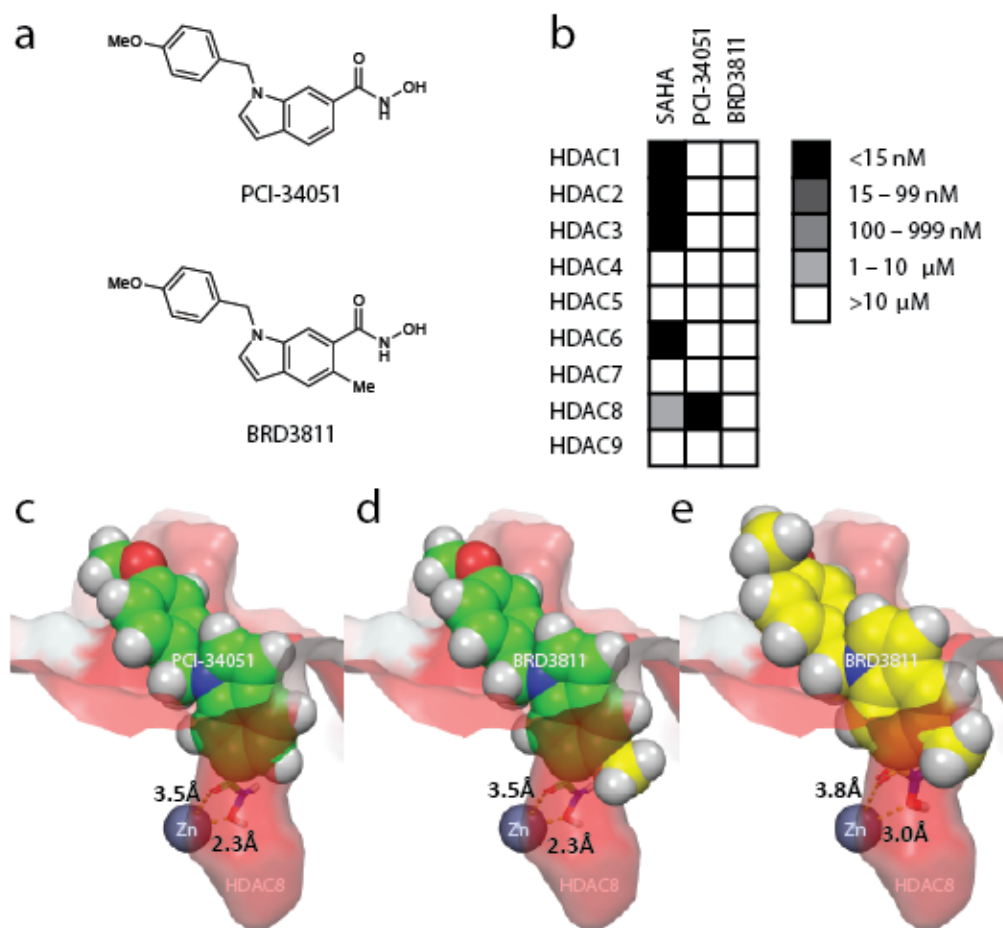
<sup>12</sup> Chapter was edited and includes a longer discussion of my *in vitro* peptide experiments.

<sup>13</sup> I performed all of the experimental studies with the peptide substrates (Figure 4.3 and 4.4) and I wrote all of the text pertaining to these experiments.

have become important therapeutic targets for a number of disease states including cancer[6, 7] and psychiatric illnesses[8]. HDACs deacetylate a large number of non-histone proteins[2, 9], and identification of the endogenous substrates of HDAC enzymes is a fundamental area of HDAC research. Of all the HDACs, HDAC8 is the best characterized structurally and kinetically[10-23]. As a result, HDAC8 can be used as a platform to develop methods for the identification of substrate specificity determinants in HDAC isozymes. Despite extensive studies, only two cellular HDAC8 substrates have been identified[24], namely, the estrogen-related receptor- $\alpha$ [25] (ERR- $\alpha$ ) and the structural maintenance of chromosome 3[26] (SMC3) protein. The complete HDAC8 substrate set remains undetermined, which is largely due to difficulties in distinguishing binding partners from substrates (Chapter 1).

To elucidate the cellular substrates and better define the biology of HDAC8, we undertook an unbiased, chemical biology approach that involves monitoring global acetylation and gene expression changes in a representative cell line following treatment with a known, potent, and highly selective HDAC8 small molecule inhibitor. Small molecule modulation coupled with mass spectrometry can offer distinct advantages for the identification of acetylation substrates and specific acetyl lysine sites responsive to HDAC8 in comparison to protein knockdown, knockout, or pulldown approaches, including: 1) de-convolution of catalytic vs. scaffolding functions associated with HDACs[27], 2) temporal control, 3) increased resolution and sensitivity, and 4) the avoidance of complications associated with transient and/or metastable interactions and complexes. Therefore, we focused on using the highly selective and potent HDAC8 inhibitor, PCI-34051[28] as well as a suitably designed negative control compound to account for potential compound driven off-target effects (Figure 4.1a,b). The inclusion of a negative control compound was particularly important, as PCI-34051 contains a metal-chelating hydroxamic acid group known to bind a variety of metalloenzymes[29]. As such, we designed and synthesized BRD3811 (Figure 4.1a), a compound that retains the hydroxamic acid functionality but contains a minor structural modification to PCI-34051 (i.e., a single methyl group introduced ortho to the hydroxamic acid group) resulting in over 1,000-fold reduction in potency for HDAC8 inhibition (Figure 4.1b). Consistent with this finding, molecular docking of PCI-34051 and BRD3811 into the active site of the HDAC8 crystal structure 2v5w reveals that

the methyl group of BRD3811 cannot be accommodated in the catalytic binding domain of HDAC8 while maintaining an optimal zinc chelation geometry (Figure 4.1c).



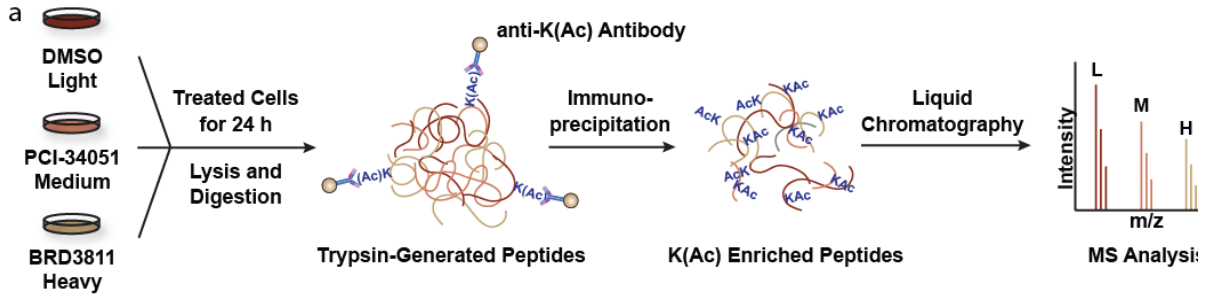
**Figure 4.1 Chemical tools for studying HDAC8**

(a) Chemical structures of the HDAC8 inhibitor PCI-34051 and the structurally related negative control compound BRD3811. (b) HDAC inhibitor potencies for PCI-34051, BRD3811, and the pan-inhibitor SAHA. (c) PCI-34051 docked into the 2V5W crystal structure of HDAC8. (d) Replacement of the ortho-hydrogen in the docked structure of PCI-34051 (c) with a methyl group. The methyl group protrudes from the enzyme pocket. (e) BRD3811 docked into the 2V5W crystal structure of HDAC8. Again, the methyl group protrudes from the enzyme pocket.

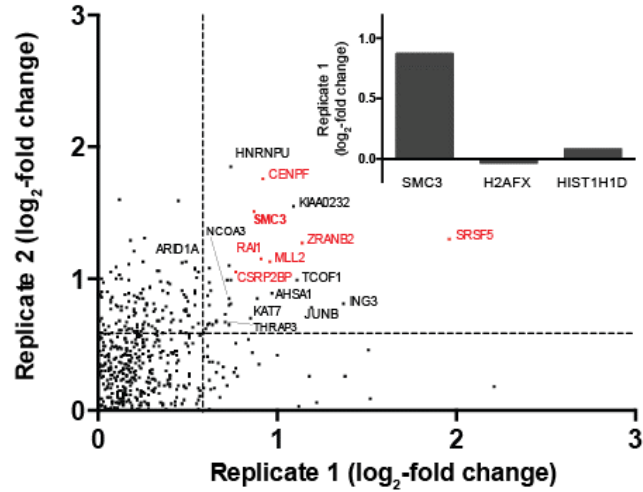
Using these chemical tools, we compared the changes in global acetylation in a representative cell line known to express HDAC8[30] (i.e. MCF7) after treatment with each compound using Stable Isotope Labeling of Amino Acids in Cell Culture (SILAC)-based quantitative mass spectrometry (MS). Briefly, cells were grown in the presence of light, medium, or heavy arginine

and lysine then treated with either PCI-34051 (10  $\mu$ M), BRD3811 (10  $\mu$ M), or vehicle (DMSO) for 24 hours (Figure 4.2a). Global acetylation profiling was completed by digesting cellular proteins with trypsin and enriching acetylated peptides by immunoprecipitation using an antibody specific for acetylated lysine residues. Over two replicates, 1,361 acetylation sites were quantified using high-resolution MS.

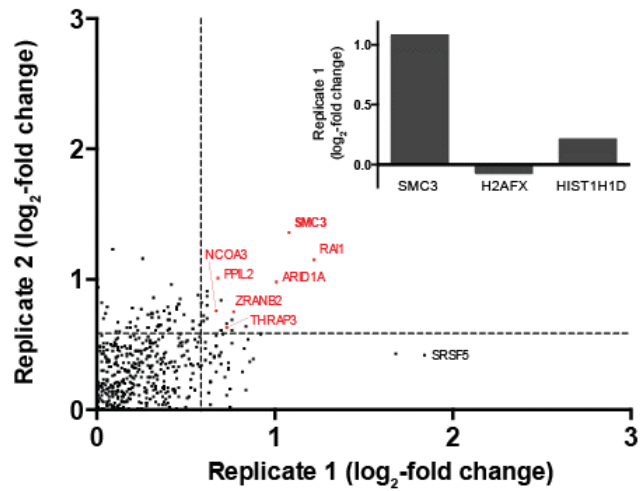




**b** Acetylation Sites PCI-34051 vs DMSO



**c** Acetylation Sites PCI-34051 vs BRD3811



**d** PCI-34051 vs DMSO ( $P \leq 0.05$ )

Protein	Position of Ac	Localization	Function
SMC3	K106	Nucleus	Mitosis
RAI1	K1087	Nucleus & Cytoplasm	Transcription
ZRANB2	K54	Nucleus	Splicing
SRSF5	K167	Nucleus	Splicing
CSRP2BP	K292	Nucleus & Cytoplasm	Acetyltransferase
MLL2	K3579	Nucleus	Methyltransferase
CENPF	K1341	Nucleus & Cytoplasm	Mitosis

**e** PCI-34051 vs BRD3811 ( $P \leq 0.05$ )

Protein	Position of Ac	Localization	Function
SMC3	K106	Nucleus	Mitosis
RAI1	K1087	Nucleus & Cytoplasm	Transcription
ZRANB2	K54	Nucleus	Splicing
ARID1A	K1808	Nucleus	Chromatin Remodeling
THRAP3	K387	Nucleus	Splicing
NCOA3	K687	Nucleus & Cytoplasm	Transcription
PPIL2	K482	Nucleus	Protein Folding

**Figure 4.2 Scheme and results of mass spectrometry identification of HDAC8 substrates**

(a) Schematic of experimental design. (b) Acetylated proteins regulated by treatment with PCI-34051 as compared to DMSO or (c) BRD3811 as the control. Each axis represents a single replicate and denotes log<sub>2</sub>-fold changes in acetylation with dashed lines indicating a 1.5-fold change in acetylation. Proteins that passed a p-value cutoff of  $\leq 0.05$  in both replicates and were not regulated by the negative control compound BRD3811 are highlighted in red. Insets show log<sub>2</sub>-fold changes in acetylation for select examples of replicate 1 and emphasize the relative lack of histone acetylation by comparison to SMC3 for each experiment. (d) Tables depicting acetylation sites regulated by more than 1.5-fold and passing p-value cutoffs of  $\leq 0.05$  in both replicates when PCI-34051 treatment was compared to DMSO or to (e) BRD3811 treatment as the control. These proteins are the same as those highlighted in red above, their subcellular localization is predominantly nuclear, and they are involved in a variety of cellular processes including gene transcription and RNA splicing.

This approach enabled us to identify numerous sites in proteins where acetylation increased by more than 1.5-fold in each of two replicates upon treatment with PCI-34051 relative to DMSO (Figure 4.2b). Of these, 7 passed a p-value cutoff of  $\leq 0.05$  (Figure 4.2b, red; Figure 4.2d) and were not regulated upon treatment with negative control compound BRD3811 relative to DMSO (Figure 4.2b,c). Alternatively, a direct comparison of PCI-34051 treatment to BRD3811 treatment (Figure 4.2c) revealed 20 proteins whose acetylation increased by more than 1.5-fold with 7 passing a p-value cutoff of  $\leq 0.05$  in two replicates (Figure 4.2c, red; Figure 4.2e). From these two data sets, we deemed 5 proteins (i.e., SMC3, RAI1, ZRANB2, NCOA3, and THRAP3) to be high-confidence substrates for HDAC8 as they were identified by comparing PCI-34051 treatment to both DMSO as well as to the negative control compound. Furthermore, ARID1A and SRSF5 were also considered candidate substrates for HDAC8 as they narrowly fell outside the bounds of our arbitrary cutoffs (i.e., 1.5-fold change and p-value  $\leq 0.05$ ) in only one of four replicates.

Importantly this unbiased approach successfully identified SMC3, a known substrate of HDAC8[26], as being significantly regulated by treatment with the HDAC8-selective inhibitor and not BRD3811. Furthermore, we were able to demonstrate for the first time that this acetylation occurs specifically on K106 of SMC3. Our coverage of acetylated proteins did not include ERR- $\alpha$ ; and therefore, we cannot verify its regulation by HDAC8 in MCF7 cells. Our coverage did include several histone loci (Figure 4.3), and we did not observe any significant changes in histone acetylation status upon treatment with PCI-34051 when compared to DMSO or BRD3811.

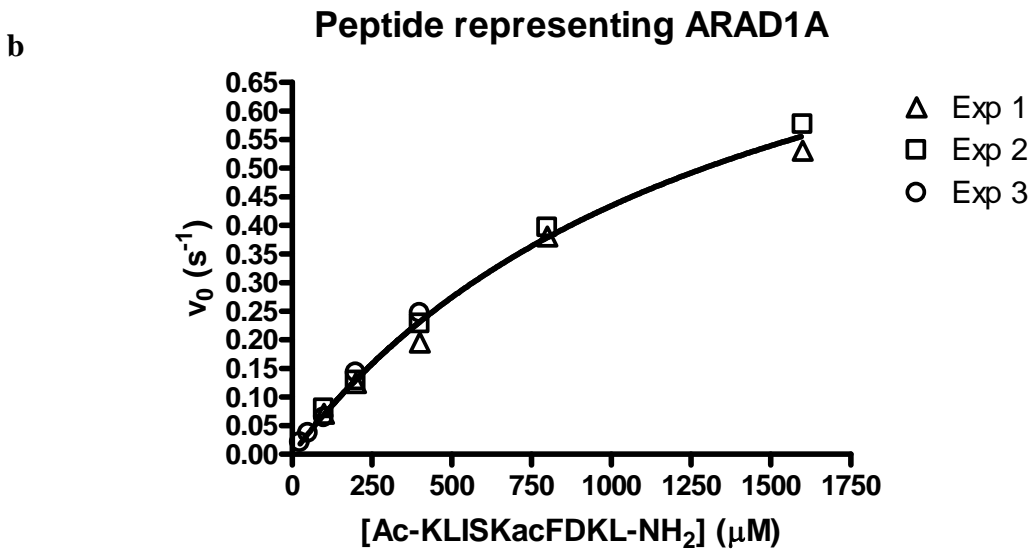
To further validate the acetylated proteins identified by our proteomics experiments as HDAC8 substrates, we assayed peptides with sequences corresponding to the identified acetylation sites. Short peptides (8 to 10 amino acids) (Figure 4.3) containing an acetylated lysine with an acetylated N-terminus and C-terminal amine were purchased. The steady state kinetic parameters were determined from a fit of the Michaelis-Menten equation to the dependence of HDAC catalyzed deacetylation on the peptide concentration. Zn(II)-HDAC8 catalyzed deacetylation of all of these peptides *in vitro* (Figure 4.3) with catalytic efficiencies varying from  $4.6 \pm 0.10$  to

740 ± 36. The peptide corresponding to ARID1A is one of the fastest peptides representing an *in vivo* protein sequence assayed thus far (Appendix A) with a  $k_{cat}/K_M$  value of 740 M<sup>-1</sup>s<sup>-1</sup> [31] (Figure 4.3). Consistent with previous work, peptides containing an aromatic residue at the +1 position (ARID1A and CSRP2BP) are the most efficient HDAC8 substrates[23]. These amino acids likely form a pi-pi bond with Tyr100 of HDAC8 to increase catalytic efficiency[24].

**a**

Peptide Deacetylation Assay (in vitro)

Protein	Ac-Sequence-NH <sub>2</sub>	$k_{cat}/K_M$ (M <sup>-1</sup> s <sup>-1</sup> )	$k_{cat}$ (s <sup>-1</sup> )	$K_M$ (μM)
RAI1	KLGGKacQRAA	11 ± 1.4	> 0.03	> 3000
ZRANB2	TEIGKacTLAEK	4.1 ± 0.30	> 0.018	> 4500
THRAP3	LGDGKacMKS	4.6 ± 0.10	> 0.015	> 3000
NCOA3	KRILHKacLLQN	50 ± 5.0	> 0.1	> 1600
SRSF5	KLSGKacEING	9.8 ± 1.9	0.018 ± 0.0060	1800 ± 900
ARID1A	KLISKacFDKL	740 ± 36	1.0 ± 0.07	1400 ± 150
CSRBP2BP	STPVKacFISR	160 ± 27	0.16 ± 0.040	970 ± 430
MLL2	SKIQKacQLDQ	32 ± 5.7	> 0.015	> 400



**Figure 4.3 Deacetylation of peptides identified by mass spectrometry**

(a) Acetylated peptides of 8-10 residues in length corresponding to the sequences surrounding the acetylation sites of a subset of identified HDAC8 substrates were subjected to an *in vitro* deacetylation assay using Zn-bound HDAC8 enzyme (see methods). Michaelis-Menten parameters for the deacetylation of each peptide are reported with the synthetic peptide corresponding to ARID1A exhibiting the largest  $k_{cat}/K_M$  reported to date for the HDAC8-catalyzed deacetylation of a non-fluorophore conjugated synthetic peptide. (b) Plot for the determination of Michaelis-Menten parameters for the ARID1A peptide.

To examine the specificity of HDAC8 reactivity with these peptides, we measured  $k_{cat}/K_M$  values for deacetylation catalyzed by the metal dependent deacetylases HDAC1-4, 6-9 (Figure 4.4). These data suggest that HDAC3 is a promiscuous enzyme, consistent with previous library studies[17, 18]. This isozyme has an aspartic acid substituted for the tyrosine in HDAC8 that interacts with the aromatic amino acid at the +1 position suggesting a change in the molecular recognition determinants. All of the acetylated peptides are substrates for HDAC8 with varying efficiencies and also substrates for other isozymes. It is currently unclear what the value of  $k_{cat}/K_M$  must be to identify the corresponding protein as a biological substrates. Additionally identification of known substrates and peptides are needed to address this question.

Changes between the reconstituted Zn(II)-HDAC8 and commercial HDAC8 were observed. These changes may be the result of the utilization of BSA and Triton-100 which increases reaction rate by 2- to 3-fold (Lopez and Fierke unpublished).

Protein	Ac-Sequence-NH <sub>2</sub>	HDAC1	HDAC2	HDAC3	HDAC4	HDAC6	HDAC7	HDAC8	HDAC9
RAI1	KLGGKacQRAA	50 ± 42	< 5	1700	-	-	-	< 5	< 5
ZRANB2	TEIGKacTLAEK	< 5	< 5	53 ± 32	8.9 ± 12	270	23 ± 5.1	< 5	< 5
SRSF5	KLSGKacEING	< 5	< 5	55 ± 18	72 ± 5.7	820	-	< 5	< 5
ARID1A	KLISKacFDKL	53 ± 17	20 ± 3.1	2500	< 5	1200	< 5	2400	< 5
THRAP3	LGDGKacMKS	< 5	34 ± 16	1600	-	-	-	< 5	< 5
NCOA3	KRILHKacLLQN	73 ± 6.2	< 5	2200	< 5	74 ± 5.0	140 ± 33	-	46 ± 36
MLL2	SKIQKacQLDQ	< 5	29.4 ± 15	220 ± 0.44	-	-	-	< 5	< 5
CSRP2BP	STPVKacFISR	51 ± 10	< 5	1500	< 5	210	-	740	< 5

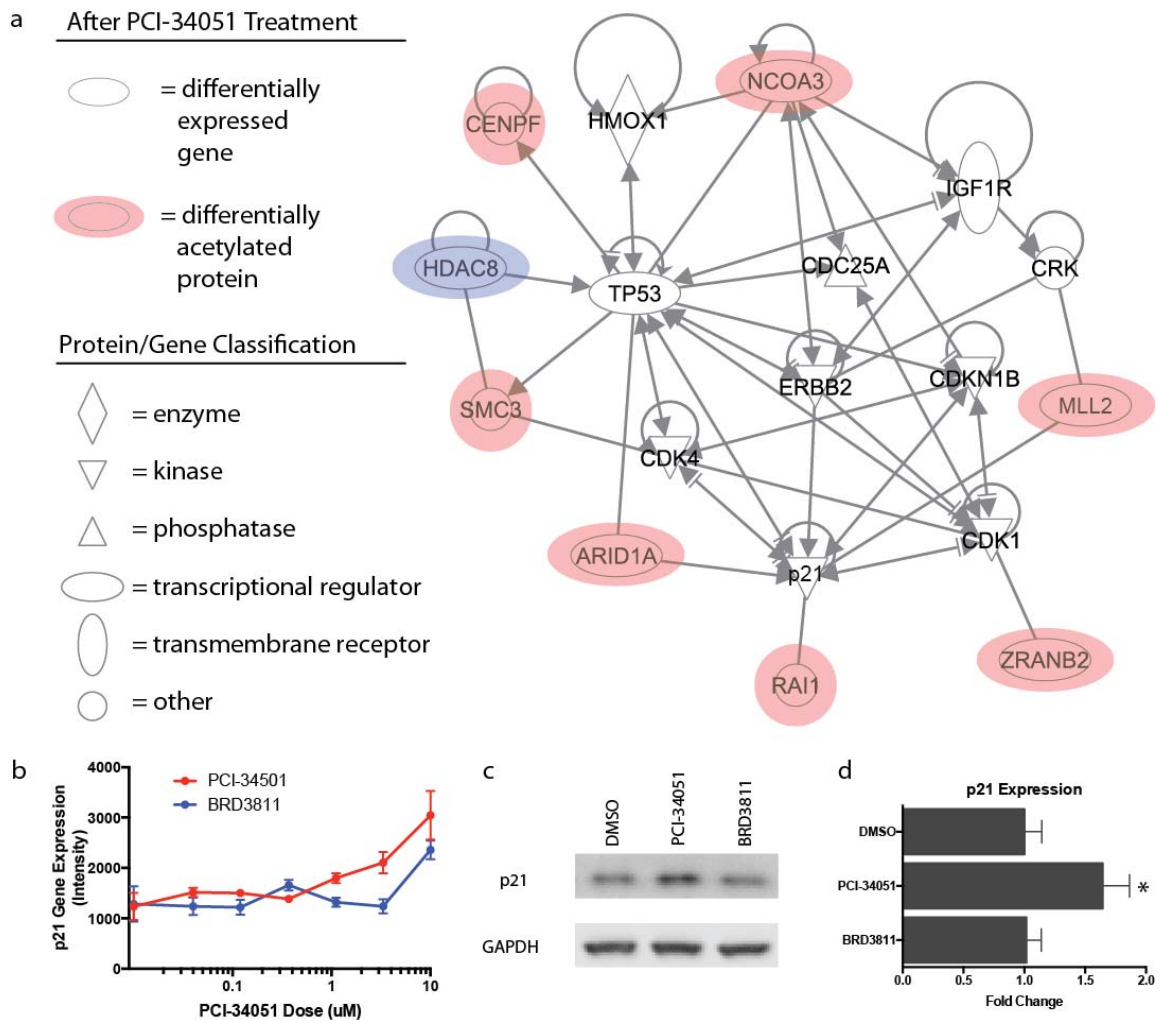
**Figure 4.4 *In vitro* peptide deacetylation using commercially available HDACs**

$k_{cat}/K_M$  ( $M^{-1}s^{-1}$ ) values for commercially available HDAC enzymes are reported as the mean of one experiment ± the standard error of the fit for 3 points. Dashes indicate particular combinations of peptides and enzyme isoforms that were not measured. Values for the metal reconstituted Zn(II)-HDAC8 enzyme (average of 3 experiments) are listed in grey for comparison.

The vast majority of the substrates identified in our study are localized in the nucleus (Figure 4.2d, e) and include transcription factors as well as proteins intimately involved in epigenetic regulation, chromatin remodeling, or RNA splicing. Interestingly, histone proteins were not

identified as substrates in our acetylome profiling experiments, and this fact was later confirmed in separate targeted SILAC experiments designed to specifically monitor histone acetylation changes. Intrigued by the non-histone but primarily nuclear nature of the candidate substrates, we tested if HDAC8 inhibition could lead to changes in gene expression independent of changes in histone acetylation. To this end, we measured the expression changes in MCF7 cells of approximately 1,000 landmark genes (L1000) as a representative measure of genome-wide effects upon treatment with PCI-34051 or BRD3811 across the dose range of 0.04 – 10  $\mu$ M (see methods). We then selected the dose-responsive genes using the IsoGene package (<http://CRAN.R-project.org/package=IsoGene>). As expected, PCI-34051 altered the expression of significantly more genes (70 genes) than did BRD3811 (6 genes).

In order to identify what non-histone proteins might be involved in PCI-34051-mediated changes in gene expression, we performed pathway analysis using Ingenuity® (<http://www.ingenuity.com/>) (Figure 4.5a) with our high confidence HDAC8 substrates (Figure 4.2d,e) and the 70 genes dose-responsive to PCI-34051 treatment (Figure 4.5) serving as our inputs. Not surprisingly, Ingenuity® did not report any direct associations between HDAC8 and the newly identified substrates; however, several of these proteins were directly linked to genes upregulated upon PCI-34051 treatment. Furthermore, we were intrigued to find that 3 of the newly identified substrates (i.e., ARID1A, RAI1, and MLL2) were directly linked to the cell cycle regulator p21 (CDKN1A). This led us to speculate that the well-known ability of ARID1A to regulate p21[32, 33] might be linked to HDAC8 inhibition, or alternatively, that RAI1 and/or MLL2 are involved. Many HDAC inhibitors are known to cause the upregulation of p21, but until now, the exact substrates responsible for mediating that effect have remained obscure[34]. When MCF7 cells are treated with PCI-34051 over the dose range 0.04 – 10  $\mu$ M, a robust dose-dependent increase in the level of p21 transcript is observed (Figure 4.5). Conversely, BRD3811 does not elicit as robust a dose-response, even though it does seem to increase the level of p21 transcript at the highest dose (i.e., 10  $\mu$ M) (Figure 4.5b). However, when assayed at the protein level, PCI-34051 treatment increased p21 (visualized via western blot) while BRD3811 treatment did not (Figure 4.5c,d).



**Figure 4.5 The tumor suppressor p21 is linked to HDAC8 inhibition**

(a) Map of the proteins identified by mass spec approach and interactions of those proteins. (b) Treatment with PCI-34051 for 24 hours results in a dose-dependent increase in p21 expression. (c) Representative western blot after treatment of MCF7 cells with either PCI-34051 or BRD3811 at 10  $\mu$ M for 48 hours shows that PCI-34051, but not the negative control, induces an increase in p21 protein levels. (d) Quantitation of western blot data from 4 independent experiments. The star denotes  $p \leq 0.05$ , relative to DMSO, as determined by a one-way analysis of variance (ANOVA) utilizing a post-hoc Dunnett's multiple comparison's test.

To further validate our findings beyond the context of a single cell type and incorporate into our analysis orthogonal biological perturbations, we expanded our gene expression studies using



L1000 into several cell lines representative of distinct tissue types: PC3 (prostate), HEPG2 (liver), HCC515 (lung), HA1E (kidney), A375 (skin), A549 (lung), HT29 (colon). We then created a gene expression signature using the 1,000 landmark genes, and compared PCI-34051 and BRD3811 treatments to the signatures of other bioactive perturbing agents using the connectivity map (cmap) database ([www.broadinstitute.org/cmap/](http://www.broadinstitute.org/cmap/)) as previously described (see methods). We integrated the results of multiple independent cmap queries using the cell lines highlighted above, and we observed that PCI-34051 treatment was highly correlated with the overexpression of p21 across multiple cell lines while BRD3811 treatment was not. In fact, p21 overexpression was the most highly correlated overexpression perturbation with PCI-34051 treatment, with an average rank in the 96<sup>th</sup> percentile (BRD3811 treatment did not correlate well with p21 overexpression, 66<sup>th</sup> percentile rank). It is quite attractive to speculate that the anti-cancer effects of the HDAC8 inhibitor PCI-34051[35] might be mediated by increasing p21 levels through these newly discovered substrates.

In conclusion, we have identified several novel substrates of HDAC8 by taking an unbiased approach coupling chemical tools with acetylome profiling. The proteins identified include the known HDAC8 substrate SMC3, but do not include histones. Furthermore, these candidate substrates were predominantly nuclear and involved in a diverse range of cellular functions including transcription and RNA splicing. We demonstrated through *in vitro* enzymatic assays as well as gene and protein expression studies that inhibition of HDAC8 can affect acetylation status and influence the levels of downstream proteins. Our experimental design relied on using BRD3811, a negative control compound based on the structure of the potent and selective HDAC8 inhibitor PCI-34051. Our approach represents a general strategy that should prove useful in future studies aimed at the identification of bona fide HDAC substrates or in elucidating the off-target effects of HDAC inhibitors.

## **Acknowledgments**

We thank Florence Wagner, Michel Weïwer, and Andrea Desouza for assistance and helpful discussions. This work was supported by the Stanley Medical Research Institute, the LINCS

Program (Grant U54 HG006093 Large scale gene expression analysis of cellular states), the National Institutes of Health (Grants NIGMS GM40602 (CAF), T32-GM-008353 (NAW), and T32-GM-008597 (CAP)), and the National Science Foundation (graduate fellowship DGE0718128 (NAW)). Steve Carr for funding. We also thank Leslie Gaffney for her help with illustrations.

## **Methods**

### **Reagents**

ATP, Coenzyme A, NAD<sup>+</sup>, L-malic acid, citrate synthase (CS), and malate dehydrogenase (MDH) were purchased from Sigma. Fluor de Lys peptide and the developing reagent were purchased from Enzo Life Sciences. The peptides were purchased from Peptide2.0. Cobalt(II) was purchased as an ICP standard from GFS Chemicals and the acetic acid standard was purchased from Ricca Chemical Company. Chelex 100 resin was purchased from Bio-Rad. All HDACs (with the exception of HDAC8 expressed and purified in our lab) were purchased from BP Biosciences and given to us as a generous gift from Edward Holson. All other materials were purchased from Fisher and were of a purity >95 % unless otherwise noted.

### **HDAC8 expression and purification**

A freshly transformed plate of BL21(DE3)pHD4-HDAC8-TEV-His<sub>6x</sub> or BL21(DE3)pHD4-HDAC8(codon optimized)-TEV-His<sub>6x</sub> was used to inoculate 5 mL of 2xYT supplemented with 50 µg/mL ampicillin and grown at 37°C. Upon reaching an OD<sub>600</sub> ~1 the cells were transferred to a flask containing 12.5 mL of 2xYT (per liter of expression media) containing 50 µg/mL ampicillin. These cells were grown to OD<sub>600</sub> > 0.6 and the 12.5 mL 2xYT was used to inoculate 1 L of modified-autoinduction TB medium (12 g/L tryptone, 24 g/L yeast extract, 8.3g/L Tris-HCl, 4 g/L lactose, 1 g/L glucose, 10 mL/L glycerol, pH 7.4) supplemented with 50 µg/mL ampicillin and 200 µM ZnSO<sub>4</sub>. The cells were grown overnight at 30°C with shaking at ~170 rpm and harvested 20-24 hours post inoculation by centrifugation (9,000 x g, 10 min, 4°C) and

then either lysed immediately or frozen as a dry cell pellet at  $-80^{\circ}\text{C}$ . Before lysing, cells were resuspended in low salt DEAE buffer (50 mM HEPES, 200  $\mu\text{M}$   $\text{ZnSO}_4$ , 1 mM TCEP, 50 mM NaCl, 5 mM KCl, 1  $\mu\text{g}/\text{mL}$  TAME, 10  $\mu\text{g}/\text{mL}$  PMSF, pH 7.8) and passed through a microfluidizer (Microfluidics) twice. The lysate was incubated with 0.1% PEI (polyethylenimine) (pH 7.9) for 10-15 minutes stirring on ice. The lysate was pelleted by centrifugation (16,800 x g, 45 min,  $4^{\circ}\text{C}$ ) and the supernatant was run through a 100 mL DEAE Sepharose column pre-equilibrated with low salt DEAE buffer (50 mM HEPES, 200  $\mu\text{M}$   $\text{ZnSO}_4$ , 1 mM TCEP, 100 mM NaCl, 5 mM KCl, pH 7.8). The column was washed with 300 mL low salt DEAE buffer and eluted with 200 mL high salt DEAE buffer (50 mM HEPES, 200  $\mu\text{M}$   $\text{ZnSO}_4$ , 1 mM TCEP, 500 mM NaCl, 5 mM KCl, pH 7.8). The elution was dialyzed (Spectra/Por 8K MWCO) against 50 mM HEPES, 50 mM NaCl, pH 7.8 for 1 hour at  $4^{\circ}\text{C}$  to remove the salt and  $\text{ZnSO}_4$ . The eluate was then mixed with 20 mL of chelating sepharose resin charged with  $\text{NiSO}_4$  in low imidazole buffer (50 mM HEPES, 100 mM NaCl, 25 mM imidazole, pH 7.8) and was incubated for 20-30 minutes stirring on ice. The slurry was poured into a column. The column was washed with 25 mM and 50 mM imidazole and HDAC8 was eluted with 250 mM imidazole. 0.5 mg of TEV protease per 3 mg HDAC8 was added to the eluate, and the mixture was dialyzed (Spectra/Por 8K MWCO) overnight against low imidazole buffer at  $4^{\circ}\text{C}$ . The next day, the eluate was applied to a 5 mL column of clean chelating sepharose resin charged with  $\text{NiSO}_4$  and equilibrated with low imidazole buffer. The column was washed with low imidazole buffer with 50 mM. The flow through and wash were concentrated to  $<3$  mL using 30,000 MWCO Amicon Ultra-15 centrifugal units and loaded onto a GE HiPrep 16/60 Sephacryl S200 HR size exclusion column equilibrated with size exclusion buffer (30 mM HEPES, 150 mM NaCl, 1 mM TCEP, pH 8). The fractions of  $>95\%$  pure HDAC8 (confirmed by SDS page) were combined and concentrated to  $<0.7$  mL in 30,000 MWCO Amicon Ultra-15 centrifugal units. The HDAC8 was dialyzed against 4 L of metal free chelation buffer (30 mM MOPS, 3 mM KCl, 1 mM EDTA, 1 mM TCEP) overnight at  $4^{\circ}\text{C}$  followed by metal free dialysis buffer (30 mM MOPS, 3 mM KCl, 1 mM EDTA, 1 mM TCEP). Finally, HDAC8 was applied to a PD-10 column in metal free dialysis buffer on a PD-10 column. All the fractions containing HDAC8 were concentrated  $>500$   $\mu\text{M}$ , flash frozen in 10  $\mu\text{L}$  aliquots, and stored at  $-80^{\circ}\text{C}$ . HDAC8 activity was confirmed using the Fluor de Lys assay as described previously[21]. This protocol utilizing pHD4-HDAC8-TEV-His<sub>6x</sub> yielded  $\sim 1.5$  mg/L HDAC8 whereas BL21(DE3)pHD4-HDAC8(codon optimized)-TEV-

His<sub>6x</sub> yielded ~3mg/L.

## Peptides

The lengths of these peptides were chosen to include an aromatic amino acid or lysine for the measurement of peptide concentration while maintaining a length of 8 to 10 amino acids that is equivalent for all peptides. The peptides solubilized in water, 50 % acetonitrile, or 10 % DMSO as needed. Once solubilized the pep20 set of peptides were chelated with Chelex resin. Peptide concentrations were measured using the fluorescamine assay when possible, or ND-1000 (NanoDrop), or just assumed based on the molecular weight and mass of the dried peptide. Peptide (0 – 800  $\mu$ M) were incubated in HDAC8 assay buffer (50 mM HEPES, 137 mM NaCl, 3.7 mM KCl, pH 7.8) and the reactions were initiated by adding 1  $\mu$ M (final concentration) Zn(II)-HDAC8 and incubated at 30°C for 0 to 120 minutes.

## Peptide deacetylation assay

Peptides were assayed using coupled assay system which enzymatically links the formation of acetate to the fluorogenic molecule NADH[23]. The reactions quenched by addition of 0.37 % (v/v, final concentration), flash frozen, and stored at -80°C. Upon thawing, the reactions were neutralized by addition of 0.6% (w/v, final concentration) NaHCO<sub>3</sub> and a coupler mixture (50 mM HEPES, 400  $\mu$ M ATP, 10  $\mu$ M NAD<sup>+</sup>, 30  $\mu$ M CoA, 0.07 U/ $\mu$ L citrate synthase, 0.04 U/ $\mu$ L malate dehydrogenase, 50  $\mu$ M acetyl CoA synthetase, 100 mM NaCl, 3 mM KCl, 50 mM MgCl, 2.5 mM L-malic acid, pH 8) was added at a ratio of 10  $\mu$ L/65  $\mu$ L and incubated for 40 minutes at room temperature in a 96 well black plate. Then NADH fluorescence (ex. = 340 nm, em. = 460 nm) was measured. Linear portions of the assay data were fit to a line yielding an initial reaction velocity.

## Fluorescamine assay

Standard curves using the peptide Ac-KGGAKacW-NH<sub>2</sub> (50 – 200 μM in water) and the fluorescamine reagent were prepared. The absorbance at 280 nm of these standards were measured using the NanoDrop spectrophotometer (Thermo Fisher Scientific) and the concentrations were calculated using the extinction coefficient 5690 μM<sup>-1</sup> cm<sup>-1</sup>[36]. For the fluorescamine standard curve the Ac-KGGAKacW-NH<sub>2</sub> peptide mixed with 0.75 M borate (final concentration, pH 9) and then 0.13 mM fluorescamine (final concentration) in a 96 well plate, (corning #3686) incubated at room temperature for 10 minutes, and the fluorescence (ex. 340 nm em. 460 nm) was measured. The peptide concentrations were calculated utilizing the Ac-KGGAKacW-NH<sub>2</sub> standard curve.

## Compound Synthesis

PCI-34051 is commercially available. BRD3811 was synthesized according to the following procedure: A solution of methyl 5-methyl-1H-indole-6-carboxylate (100 mg, 0.529 mmol, 1.0 equiv), potassium iodide (8.8 mg, 0.053 mmol, 0.1 equiv), and sodium hydride (60 % dispersion, 23.3 mg, 0.581 mmol, 1.1 equivalents) in DMF (1.6 mL) was stirred at 0 °C for 1 hour. Next, a solution of 1-(chloromethyl)-4-methoxybenzene (124 mg, 0.793 mmol, 1.5 equivalents) in DMF (1 mL) was added, and the reaction mixture was heated to 65°C for 2 hours. The reaction was diluted in saturated with Na<sub>2</sub>CO<sub>3(aq)</sub> (50 mL) and extracted with EtOAc (3 x 25 mL). The combined organic extracts were washed with brine, dried over MgSO<sub>4</sub>, filtered, and concentrated under reduced pressure. The resulting residue was purified by flash chromatography (hexanes/EtOAc) to yield methyl 1-(4-methoxybenzyl)-5-methyl-1H-indole-6-carboxylate (99.0 mg, 0.320 mmol, 61 %). To a solution of methyl 1-(4-methoxybenzyl)-5-methyl-1H-indole-6-carboxylate (99.0 mg, 0.320 mmol, 1.0 equivalents) and sodium hydroxide (64.0 mg, 1.60 mmol, 5.0 equivalents) in 1:1 MeOH:THF (1.3 mL) was added 50 % aqueous hydroxylamine (0.628 mL, 10.2 mmol, 32 equivalents), and the resulting solution was stirred for 7 hours. Upon reaction completion, 5 mL of water was added followed by removal of organic solvents under reduced pressure. Neutralization of the remaining aqueous solution with 1 M HCl<sub>(aq)</sub> resulted in the

precipitation of product, which was filtered, washed with cold water, and dried under reduced pressure to yield BRD3811 (45.0 mg, 0.145 mmol, 45 %) as a white powder in >95 % purity (as determined by LCMS). <sup>1</sup>H NMR (300 MHz, DMSO-d<sub>6</sub>): δ 10.72 (br s, 1H), 8.97 (br s, 1H), 7.51 (d, J = 2.8 Hz, 1H), 7.41 (s, 1H), 7.37 (s, 1H), 7.16 (d, J = 8.4 Hz, 2H), 6.86 (d, J = 8.4 Hz, 2H), 6.40 (d, J = 2.8 Hz, 1H), 5.31 (s, 2H), 3.70 (s, 3H), 2.36 (s, 3H) ppm.

### **Caliper Assay**

The biochemical determination of HDAC IC<sub>50</sub>s for compounds was performed as described previously.(US Patent Application WO 2013067391) All HDACs were purchased from BPS Bioscience. The substrates Broad Substrate A and Broad Substrate B were synthesized in house, but can be purchased from PerkinElmer (Product number CLS960006 and CLS960007, respectively). All other reagents were purchased from Sigma. Caliper EZ reader II system was used to collect all data. Compounds were tested in duplicate in a 12-point dose curve with 3-fold serial dilution starting from 33.33 μM. Purified HDACs were incubated with 2 μM carboxyfluorescein (FAM)-labeled acetylated or trifluoroacetylated peptide substrate (Broad Substrate A and B, respectively) and test compound for 60 min at room temperature, in HDAC assay buffer that contained 50 mM HEPES (pH 7.4), 100 mM KCl, 0.01% BSA and 0.001% Tween-20. Reactions were terminated by the addition of the known pan HDAC inhibitor LBH-589 (panobinostat) with a final concentration of 1.5 μM. Substrate and product were separated electrophoretically and fluorescence intensity in the substrate and product peaks was determined and analyzed by Labchip EZ Reader. The percent inhibition was plotted against the compound concentration, and the IC<sub>50</sub> values were automatically fitted by Genedata Screener software using 4-parameter logistic dose response model.

### **Molecular Docking**

The best pose of compound PCI-34051 in the HDAC8 binding site was determined using the Zn-HDAC8 structure deposited in the PDB by Vannini et al. (PDBID 2v5w)[11]. Protonation of the

structure and formation of side-chain–metal interactions was performed using the Protein Preparation Wizard followed by manual correction in the Schrodinger Drug Discovery Suite 2013-2. (<http://www.schrodinger.com/smdd/>) Histidines 142 and 143 were protonated at the delta nitrogen and histidine 180, which interacts with zinc at the delta position was protonated at the epsilon position. The amido lysine residue of the native peptide ligand (L.ALY5) was used as the center for the docking grid. The final cubic docking grid had the dimensions 21x21x21Å<sup>3</sup> constructed using Schrodinger Glide 6.0 with default parameters. Glide XP was used to dock PCI-34051 and BRD3811 with options to enhance planarity of aromatic groups and post-docking minimization[37]. The pose with the highest Glide XP docking score for each is depicted in Figure 4.1 using PyMol. PCI-34051 had a Glide XP score of -6.1 and the methyl analog, BRD3811, had a score of only -3.5 with significant decreases in the hydrogen bonding and Coulombic score contributions and loss of the metal binding reward. Though even the former could be considered low in some virtual screening targets, metal binding is difficult to account for accurately. In this system it is handled approximately through limited charge interactions and a metal reward added to the score. The favorable interaction energy would likely be improved for PCI-34051 if metal interaction were correctly predicted. Conversely, the geometry of BRD3811 does not permit binding with the metal center and the score should be considered exponentially worse than PCI-34051.

### **Cell Culture and Compound Treatment for Proteomics Experiments**

MCF7 cells were grown and expanded from the same frozen vial stock, followed by differentially labeling with non-radioactive stable isotopic amino acids by growing in light, medium, and heavy SILAC media (See below for recipe), respectively. MCF7 cells grown in SILAC medium were plated (2 million cells/plate, 10 mL per plate) into 10-cm tissue culture treated plates and incubated 24 hours prior to treatment. For treatment, 10 µl of compounds (10 mM stocks in DMSO) or DMSO vehicle control were added to the plates, and the cells were incubated for 24h. Next, the growth medium was aspirated, and the monolayers of cells were rinsed twice with cold-PBS. Cells were detached using a cell scraper and collected with 1 ml of

cold PBS. Cell pellets were harvested by centrifugation at 1,500 rpm for 1 min and flash frozen in liquid nitrogen. All pellets were stored in -80 °C freezer prior to lysis.

SILAC media

450 mL DMEM

50 mL FBS (Sigma, F-0392)

5 mL 100x Pen/Strep/Glutamine (Gibco 10378)

3.9 mL 45% Glucose solution (Sigma, G8769)

500 µL Methionine (stock 30g/L, final 30mg/L)

500 µL Proline (stock 20g/L, final 20mg/L)

500 µL Lysine (stock 146g/L, final 146mg/L)

500 µL Arginine (stock 84g/L, final 84mg/L)

## **Proteomics**

SILAC-labeled MCF-7 cells were lysed in ice-cold 8 M urea, 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 2 ug/ml aprotinin (Sigma-Aldrich), 10 ug/ml leupeptin (Roche Applied Science), 1 mM phenylmethylsulfonyl fluoride (PMSF), and 5 mM sodium butyrate (Sigma-Aldrich). Lysates were centrifuged at 20,000 x g for 10 min at 4 °C to remove insoluble material. Protein concentrations were measured using a bicinchoninic acid (BCA) protein assay. For each replicate, 10 mg of protein per SILAC state was used for acetylation profiling. Proteins were reduced with 5 mM dithiothreitol for 45 min at room temperature. After reduction, proteins were alkylated using 10 mM iodoacetamide for 30 min at room temperature in the dark. Samples were diluted to 2 M urea and digested overnight with sequencing grade trypsin (Promega) using an enzyme to substrate ratio of 1:50 (w/w). Trifluoroacetic acid (TFA) was used to quench digests.



Peptide samples were desalted on tC18 SepPak SPE cartridges (Waters) exactly as previously described[38]. Peptides were fractionated by basic pH reversed-phase (bRP) chromatography exactly as previously described[38]. Briefly, a Zorbax 300 Extend-C18 column (9.4 × 250 mm, 300 Å, 5 µm; Agilent) was used for the separation. Peptides were reconstituted in 5 mM ammonium formate (pH 10.0)/2 % (vol/vol) acetonitrile (bRP Buffer A). Using the exact method parameters previously described[38], a total of 96 2 ml fractions were collected across the bRP separation. For acetylated (Kac) peptide analysis, each fraction was combined in a non-contiguous manner such that every eighth fraction was combined (final fraction 1 = 1,9,17,25,33,41,49,57,65; final fraction 2 = 2,10,18,26,34,42,50,58,66; ...) to create 8 final fractions. Pooled fractions were dried using vacuum centrifugation. An acetyl lysine antibody (Immunchem) was used for enrichment of Kac peptides from fractionated samples. Dried samples were reconstituted in 1.5 ml of 50 mM MOPS (pH 7.2), 10 mM sodium phosphate and 50 mM NaCl (IP buffer). Peptides were incubated with 120 µg of anti-Kac antibody beads for 1 hour at 4 °C with end-over-end rotation. Antibody beads were washed twice with 1.5 ml of ice-cold IP buffer followed by three washes with ice-cold PBS. Kac peptides were eluted from the antibody with 2 x 50 µL of 0.15 % TFA. Enriched peptides were desalted using StageTips exactly as previously described[38]. Samples were analyzed by nanoflow-UPLC-HCD-MS/MS using an Easy-nLC 1000 system (Proxeon) coupled online to a Q Exactive mass spectrometer (Thermo Fisher Scientific). Samples were reconstituted in 9 µL of 3 % MeCN/0.1 % FA and 4 µL was injected for analysis. Samples were injected on at a flow rate of 500 nl/min onto a PicoFrit column (360 µm (OD) × 75 µm (ID)), 10 µm ID tip, 50 cm length (New Objective) self-packed with 24 cm of ReproSil-Pur 120 Å, 1.9 µm C18-AQ beads and heated to 50 °C using a column heater (Pheonix S&T). The gradient and flow rate settings used were as previously described[38]. The Q Exactive was operated by acquiring an MS1 scan (R = 70,000) followed by MS/MS scans on the 12 most abundant ions. For MS acquisition, ion targets of 3 × 10<sup>6</sup> and 5 × 10<sup>4</sup> ions were used for MS1 and MS2 scans, respectively. A maximum ion time of 20 ms and 120 ms was used for MS1 and MS2 scans, respectively. The HCD collision energy was set to 25. The dynamic exclusion time was set to 20 s and the peptide match and isotope exclusion functions were enabled. The MaxQuant software package (version 1.3.0.5) was used for identification and quantification of MS data. For searching, the enzyme specificity was set to trypsin, the maximum number of missed cleavages was set to 2, the precursor mass tolerance

was set to 20 ppm for the first search, and the tolerance was set to 6 ppm for the main search. Carbamidomethylation of cysteines was searched as a fixed modification and oxidation of methionines, N-terminal acetylation of proteins, acetylation of lysines were searched as a variable modifications. The minimum peptide length was set to 6, and false discovery rate for peptide, protein, and site identification was set to 1%.

### **Cell Culture and Compound Treatment for Gene Expression Experiments**

MCF7 cells (ATCC, #HTB-22) were cultured in DMEM medium (Gibco, #11995) containing 10% fetal bovine serum (Sigma, F4135) and 1x Pen Strep Glutamine (Gibco, #10378). Cells were plated into 384-well tissue culture treated plates (Corning, #3707) using a Multidrop Combi (Thermo, # 5840300) at 2,000 cells per well. Cells were incubated for 24 hours at 37°C in a humidified incubator containing 5 % CO<sub>2</sub> before treatment.

Prior to treatment, 10 mM DMSO stock solutions of compounds were diluted to multiple doses in DMSO and arrayed into a 384-well plate (Abgene #AB-1056). These 1,000x stock solutions were first diluted (100-fold) in culture medium, and then the diluted compounds were transferred to the cell culture plates using CyBi-well vario 384-well tips (another 10-fold dilution). Ultimately, all compounds were diluted 1,000-fold to their desired serial concentrations with a final DMSO concentration of 0.1%. Treated cells were incubated for 24 hours prior to lysis. Cells were lysed by partial removal of the culture media (15 µl remaining) followed by the addition of TCL lysis buffer (Qiagen #1031576) using a liquid handling system. Cell lysate plates were sealed using a plate sealer, kept at room temperature for 30 minutes, and then frozen at -80°C until L1000 gene expression profiling was performed. Detailed cell culture and treatment protocols for L1000 can be found at <http://lincscloud.org>.

## **Gene Expression**

This study utilized L1000, a high-throughput, bead-based gene expression assay in which mRNA is extracted from cultured human cells treated with various chemical or genomic perturbagens (small molecules, gene knockdowns, or gene over-expression constructs) as previously described[39]. This mRNA is reverse-transcribed into first-strand cDNA. Gene specific probes containing barcodes and universal primer sites are annealed to the first strand cDNA. The probes are ligated to form a template for PCR. The template is PCR amplified with biotinylated-universal primers. The end products are biotinylated, fixed length, barcoded amplicons. The amplicons can then be mixed with Luminex beads that contain complementary barcodes to those encoded in each of the 1000 amplified landmark genes. These 1000 landmark genes were chosen as a reduced representation of the transcriptome and account for the majority of expression variation across many cellular contexts (Subramanian, et al., manuscript in preparation). These beads are then stained with fluorescent streptavidin-phycoerythrin (SAPE) and detected in 384-well plate format on a Luminex FlexMap flow cytometry-based scanner. The resulting readout is a measure of mean fluorescent intensity (MFI) for each landmark gene. The raw expression data are log<sub>2</sub>-scaled, quantile normalized, and z-scored, such that a differential expression value is achieved for each gene in each well. In the standard L1000 protocol, each well corresponds to a different perturbagen and these differential expression values are collapsed across replicate wells to yield a differential expression signature for each perturbagen (Subramanian, et al., manuscript in preparation). The signatures of different perturbagens can then be compared to identify those that result in similar or dissimilar transcriptional responses as previously described[40, 41].

## **CMap Query Integration Analysis.**

A meta-analysis of CMap query results was performed (i.e., lists of perturbations ordered by similarity to an input gene expression signature) using a novel algorithm that is currently under development (Subramanian, et al., manuscript in preparation). Briefly, we took a set of signature-level connections and integrated them to yield a set of perturbagen-level connections ranked by consistent connectivity across cell lines within a given perturbation type. The summarization

algorithm works as follows: A query result is first grouped by cell line and perturbagen type (small molecule, consensus gene knockdown, or overexpression). The connectivity scores are then normalized by dividing by the signed mean score of each group. The scores are converted to percentile ranks within each group. For each unique perturbagen, an integrated score is computed as the weighted enrichment score of the best N (by default 4) connected cell types[42]. In addition the mean rankpoint for the best N cell types is reported. The output is a table of integrated scores and rankpoints for each unique perturbagen id. We currently require connections to exist in at least 4 cell lines before integrating.

## **Western Blotting**

MCF7 cells were treated with compounds for 48 h, at which time, lysates were collected using RIPA buffer with added protease (Roche) and phosphatase (Roche) inhibitors. Electrophoresis was performed using NuPage 4-12 % Bis-Tris gels (Invitrogen). Proteins were transferred to a nitrocellulose membrane and probed using antibodies for p21 (Cell Signaling) and GAPDH (Cell Signalling). Chemiluminescence was induced by subsequent incubation with HRP-linked secondary antibodies (GE Healthcare UK Ltd.) and treatment of the membrane with the appropriate ECL solutions (Thermo Scientific). Visualization was accomplished using a ChemiDoc MP System (Bio-Rad), and the raw data files were converted to jpegs using ImageJ (NIH).

## **Bibliography**

- [1] G.A. Khoury, R.C. Baliban, C.A. Floudas, Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database, *Scientific reports*, 1 (2011).
- [2] C. Choudhary, C. Kumar, F. Gnad, M.L. Nielsen, M. Rehman, T.C. Walther, J.V. Olsen, M. Mann, Lysine acetylation targets protein complexes and co-regulates major cellular functions, *Science*, 325 (2009) 834-840.

- [3] X.J. Yang, E. Seto, The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men, *Nat Rev Mol Cell Biol*, 9 (2008) 206-218.
- [4] G. Piperno, M. LeDizet, X.J. Chang, Microtubules containing acetylated alpha-tubulin in mammalian cells in culture, *The Journal of cell biology*, 104 (1987) 289-302.
- [5] S. Spange, T. Wagner, T. Heinzl, O.H. Kramer, Acetylation of non-histone proteins modulates cellular signalling at multiple levels, *The international journal of biochemistry & cell biology*, 41 (2009) 185-198.
- [6] G.P. Delcuve, D.H. Khan, J.R. Davie, Targeting class I histone deacetylases in cancer therapy, *Expert Opin Ther Targets*, 17 (2013) 29-41.
- [7] P. Marks, R.A. Rifkind, V.M. Richon, R. Breslow, T. Miller, W.K. Kelly, Histone deacetylases and cancer: causes and therapies, *Nature reviews. Cancer*, 1 (2001) 194-202.
- [8] J. Graff, L.H. Tsai, The potential of HDAC inhibitors as cognitive enhancers, *Annual review of pharmacology and toxicology*, 53 (2013) 311-330.
- [9] M.A. Glozak, N. Sengupta, X. Zhang, E. Seto, Acetylation and deacetylation of non-histone proteins, *Gene*, 363 (2005) 15-23.
- [10] A. Vannini, C. Volpari, G. Filocamo, E.C. Casavola, M. Brunetti, D. Renzoni, P. Chakravarty, C. Paolini, R. De Francesco, P. Gallinari, C. Steinkuhler, S. Di Marco, Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor, *Proc Natl Acad Sci U S A*, 101 (2004) 15064-15069.
- [11] A. Vannini, C. Volpari, P. Gallinari, P. Jones, M. Mattu, A. Carfi, R. De Francesco, C. Steinkuhler, S. Di Marco, Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex, *EMBO Rep*, 8 (2007) 879-884.
- [12] J.R. Somoza, R.J. Skene, B.A. Katz, C. Mol, J.D. Ho, A.J. Jennings, C. Luong, A. Arvai, J.J. Buggy, E. Chi, J. Tang, B.C. Sang, E. Verner, R. Wynands, E.M. Leahy, D.R. Dougan, G. Snell, M. Navre, M.W. Knuth, R.V. Swanson, D.E. McRee, L.W. Tari, Structural snapshots of human HDAC8 provide insights into the class I histone deacetylases, *Structure*, 12 (2004) 1325-1334.
- [13] M. Marek, S. Kannan, A.T. Hauser, M. Moraes Mourao, S. Caby, V. Cura, D.A. Stolfi, K. Schmidtkunz, J. Lancelot, L. Andrade, J.P. Renaud, G. Oliveira, W. Sippl, M. Jung, J. Cavarelli, R.J. Pierce, C. Romier, Structural basis for the inhibition of histone deacetylase 8 (HDAC8), a key epigenetic player in the blood fluke *Schistosoma mansoni*, *PLoS Pathog*, 9 (2013) e1003645.

- [14] L. Whitehead, M.R. Dobler, B. Radetich, Y. Zhu, P.W. Atadja, T. Claiborne, J.E. Grob, A. McRiner, M.R. Pancost, A. Patnaik, W. Shao, M. Shultz, R. Tichkule, R.A. Tommasi, B. Vash, P. Wang, T. Stams, Human HDAC isoform selectivity achieved via exploitation of the acetate release channel with structurally unique small molecule inhibitors, *Bioorg Med Chem*, 19 (2011) 4626-4634.
- [15] D.P. Dowling, S.L. Gantt, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structural studies of human histone deacetylase 8 and its site-specific variants complexed with substrate and inhibitors, *Biochemistry*, 47 (2008) 13554-13563.
- [16] D.P. Dowling, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structures of metal-substituted human histone deacetylase 8 provide mechanistic inferences on biological function, *Biochemistry*, 49 (2010) 5048-5056.
- [17] Z.A. Gurard-Levin, K.A. Kilian, J. Kim, K. Bahr, M. Mrksich, Peptide arrays identify isoform-selective substrates for profiling endogenous lysine deacetylase activity, *ACS Chem Biol*, 5 (2010) 863-873.
- [18] Z.A. Gurard-Levin, J. Kim, M. Mrksich, Combining mass spectrometry and peptide arrays to profile the specificities of histone deacetylases, *Chembiochem*, 10 (2009) 2159-2161.
- [19] Z.A. Gurard-Levin, M. Mrksich, The activity of HDAC8 depends on local and distal sequences of its peptide substrates, *Biochemistry*, 47 (2008) 6242-6250.
- [20] D. Riester, C. Hildmann, S. Grunewald, T. Beckers, A. Schwienhorst, Factors affecting the substrate specificity of histone deacetylases, *Biochem Biophys Res Commun*, 357 (2007) 439-445.
- [21] S.L. Gantt, S.G. Gattis, C.A. Fierke, Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion, *Biochemistry*, 45 (2006) 6170-6178.
- [22] S.L. Gantt, C.G. Joseph, C.A. Fierke, Activation and inhibition of histone deacetylase 8 by monovalent cations, *J Biol Chem*, 285 (2010) 6036-6043.
- [23] N.A. Wolfson, C.A. Pitcairn, E.D. Sullivan, C.G. Joseph, C.A. Fierke, An enzyme-coupled assay measuring acetate production for profiling histone deacetylase specificity, *Anal Biochem*, (2014).
- [24] N.A. Wolfson, C.A. Pitcairn, C.A. Fierke, HDAC8 substrates: Histones and beyond, *Biopolymers*, 99 (2013) 112-126.

- [25] B.J. Wilson, A.M. Tremblay, G. Deblois, G. Sylvain-Drolet, V. Giguere, An acetylation switch modulates the transcriptional activity of estrogen-related receptor alpha, *Mol Endocrinol*, 24 (2010) 1349-1358.
- [26] M.A. Deardorff, M. Bando, R. Nakato, E. Watrin, T. Itoh, M. Minamino, K. Saitoh, M. Komata, Y. Katou, D. Clark, K.E. Cole, E. De Baere, C. Decroos, N. Di Donato, S. Ernst, L.J. Francey, Y. Gyftodimou, K. Hirashima, M. Hullings, Y. Ishikawa, C. Jaulin, M. Kaur, T. Kiyono, P.M. Lombardi, L. Magnaghi-Jaulin, G.R. Mortier, N. Nozaki, M.B. Petersen, H. Seimiya, V.M. Siu, Y. Suzuki, K. Takagaki, J.J. Wilde, P.J. Willems, C. Prigent, G. Gillessen-Kaesbach, D.W. Christianson, F.J. Kaiser, L.G. Jackson, T. Hirota, I.D. Krantz, K. Shirahige, HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle, *Nature*, 489 (2012) 313-317.
- [27] S.H. You, H.W. Lim, Z. Sun, M. Broache, K.J. Won, M.A. Lazar, Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 in vivo, *Nature structural & molecular biology*, 20 (2013) 182-187.
- [28] S. Balasubramanian, J. Ramos, W. Luo, M. Sirisawad, E. Verner, J.J. Buggy, A novel histone deacetylase 8 (HDAC8)-specific inhibitor PCI-34051 induces apoptosis in T-cell lymphomas, *Leukemia*, 22 (2008) 1026-1034.
- [29] D.M. Griffith, B. Szocs, T. Keogh, K.Y. Suponitsky, E. Farkas, P. Buglyo, C.J. Marmion, Suberoylanilide hydroxamic acid, a potent histone deacetylase inhibitor; its X-ray crystal structure and solid state and solution studies of its Zn(II), Ni(II), Cu(II) and Fe(III) complexes, *Journal of inorganic biochemistry*, 105 (2011) 763-769.
- [30] K. Ververis, T.C. Karagiannis, An atlas of histone deacetylase expression in breast cancer: fluorescence methodology for comparative semi-quantitative analysis, *American journal of translational research*, 4 (2012) 24-43.
- [31] A.S. Madsen, C.A. Olsen, Profiling of substrates for zinc-dependent lysine deacylase enzymes: HDAC3 exhibits decrotonylase activity in vitro, *Angew Chem Int Ed Engl*, 51 (2012) 9083-9087.
- [32] B. Guan, T.L. Wang, M. Shih Ie, ARID1A, a factor that promotes formation of SWI/SNF-mediated chromatin remodeling, is a tumor suppressor in gynecologic cancers, *Cancer Res*, 71 (2011) 6718-6727.
- [33] J.N. Wu, C.W. Roberts, ARID1A mutations in cancer: another epigenetic tumor suppressor?, *Cancer discovery*, 3 (2013) 35-43.
- [34] Z. Sun, D. Feng, B. Fang, S.E. Mullican, S.H. You, H.W. Lim, L.J. Everett, C.S. Nabel, Y. Li, V. Selvakumaran, K.J. Won, M.A. Lazar, Deacetylase-independent function of HDAC3 in

transcription and metabolism requires nuclear receptor corepressor, *Molecular cell*, 52 (2013) 769-782.

[35] S. Pramana, D. Lin, P. Haldermans, Z. Shkedy, T. Verbeke, H. Göhlmann, A. De Bond, W. Talloen, L. Bijns., IsoGene: An R Package for Analyzing Dose-response Studies in Microarray Experiments *The R Journal* 2(2010) 5-12.

[36] S.C. Gill, P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, *Anal Biochem*, 182 (1989) 319-326.

[37] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *J Med Chem*, 47 (2004) 1739-1749.

[38] N.D. Udeshi, P. Mertins, T. Svinkina, S.A. Carr, Large-scale identification of ubiquitination sites by mass spectrometry, *Nature protocols*, 8 (2013) 1950-1960.

[39] D. Peck, E.D. Crawford, K.N. Ross, K. Stegmaier, T.R. Golub, J. Lamb, A method for high-throughput gene expression signature analysis, *Genome biology*, 7 (2006) R61.

[40] J. Lamb, E.D. Crawford, D. Peck, J.W. Modell, I.C. Blat, M.J. Wrobel, J. Lerner, J.P. Brunet, A. Subramanian, K.N. Ross, M. Reich, H. Hieronymus, G. Wei, S.A. Armstrong, S.J. Haggarty, P.A. Clemons, R. Wei, S.A. Carr, E.S. Lander, T.R. Golub, The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease, *Science*, 313 (2006) 1929-1935.

[41] J. Lamb, The Connectivity Map: a new tool for biomedical research, *Nature reviews. Cancer*, 7 (2007) 54-60.

[42] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, *Proc Natl Acad Sci U S A*, 102 (2005) 15545-15550.



## Chapter 5

### **Acetylated protein substrates have enhanced reactivity with HDAC8 compared to peptide substrates**

#### **Introduction**

Histone deacetylases (HDACs) are a group of 18 proteins which catalyze the removal of the acetyl moiety from previously modified acetyl lysine residues in proteins[1, 2]. These acetylation/deacetylation events alter protein properties[3] and regulate downstream cellular events[4, 5]. Regulation of acetylation is important for effective cellular signaling and cellular homeostasis. Aberrant acetylation can cause the cell to fail to perform critical tasks resulting in diseases ranging from neurological disorders[6, 7] to cancer[8, 9]. As each HDAC is theorized to have distinct substrates, identifying the specific substrate set for each HDAC isozyme may lead to a better understanding of the role of HDACs in disease prognosis and lead to better treatments of these diseases.

There are many difficulties in determining HDAC-substrate pairs. One complication is that multiple HDACs may catalyze deacetylation of the same substrate. As a result, when one deacetylase is knocked out or inhibited, another HDAC isozyme may restore deacetylation of a target protein. This overlap in specificity has largely prevented genetic knockdowns from successfully identifying HDAC substrate sets. Furthermore, interpreting genetic studies have been difficult as it is frequently unclear whether phenotypes represent a direct interaction between HDAC and a substrate or a downstream effect caused by perturbation of acetylation of another HDAC substrate. Another difficulty is the likelihood that HDAC-substrate interactions are weak and short lived in contrast with higher affinity protein binding partners that frequently

form complexes with HDACs. Therefore is it difficult to differentiate between substrates and binding partners in pulldown experiments. Due to challenging and frequently unperformed experiments, this lack of differentiation has complicated analysis even after a putative substrate is identified. To mitigate these difficulties, the HDAC field has sought to identify sequence motifs that define each isozyme's substrate specificity[10-13]. To date, the vast majority of studies have utilized short peptides to determine HDAC recognition motifs. However, the validity of mimicking the recognition of proteins using peptides has not yet been established. The role of long range interactions and secondary structural elements is still undetermined.

While a few putative HDAC8 substrates have been identified such as ERR $\alpha$ [14] and SMC3[15], the protein substrate set for HDAC8 is currently unclear. Proteomics studies of HDAC8 have identified tens of substrates/binding partners[14-21]. Of the putative HDAC8 substrates, many are difficult to purify, preventing efforts to assay reactivity of HDAC8 with these proteins *in vitro*. Further adding to these challenges are the difficulties of isolating or preparing singly acetylated proteins for kinetic studies. As a result, previous studies of HDAC8 selectivity have analyzed reactivity with non-specifically acetylated substrates. Kinetics using these substrates provide little information about substrate preference as these protein contain multiple acetylation sites which may each be deacetylated with different rates. Furthermore, these analyses are further complicated by the fact that acetyl lysines in one position may affect the deacetylation rates of another acetyl lysine residues.

Singly acetylated proteins provide an effective starting place to investigate HDAC8 preferences. HDAC8 is located both in the cytosol and the nucleus[20, 22-24] and has been suggested to catalyze the deacetylation of histones. *In vivo* H3 acetylation levels have been shown to differ modestly upon the overexpression of HDAC8 in HEK293 cells[25], and may indicate that HDAC8 deacetylation of H3 in specific gene regions is masked by the few global acetylation changes to histones[26]. Deacetylation of core histones and peptides representing H3 shows that HDAC8 can catalyze the deacetylation of H3 *in vitro*[27, 28]. Histone H3 therefore presents an interesting target to study HDAC substrate specificity using full length proteins. Furthermore, this substrate presents a framework for future study of HDAC substrate specificity.

Many HDACs occur in complexes where the protein binding partners may alter substrate recognition[29]. HDAC8 has historically been considered to act in the absence of protein

cofactors[10-13, 30-33], and thus analyzing the specificity of HDAC8 mitigates the need to purify difficult to isolate complexes. If HDAC8 forms larger protein complexes, determination of how these complexes affect substrate selectivity is crucial to understanding HDAC substrate recognition. Finally, HDAC8 is the best biochemically characterized HDAC with numerous crystal[30, 31, 34-38], kinetic[32, 33], and substrate specificity studies[10-13]. These studies provide a good platform in which to parse HDAC substrate recognition.

In this paper, we present the first detailed kinetic study of HDAC-catalyzed deacetylation of singly acetylated full length substrates. Singly acetylated lysine side chains are incorporated into H3 using unnatural amino acid incorporation[39, 40]. Using these substrates we directly compare reactivity of peptide substrates and acetyl sites in single H3 with the same primary sequences. Furthermore, we analyze the effects of histone complexes, such as the octamer, on H3 reactivity. From these studies we demonstrate that deacetylation of acetylated full length H3 (tetramer and octamer complexes) by HDAC8 is significantly faster than acetylated peptides (>10-fold  $k_{cat}/K_M$  over peptides). In H3, HDAC8 recognizes peptide tetramer substrates largely based on the six amino acids proximal to the acetyl lysine. Taken together, these results suggest that long range contacts enhance substrate recognition while short range contacts determine substrate specificity for H3 substrates.

## **Materials and Methods**

### **Reagents**

Adenosine triphosphate (ATP), Coenzyme A, Nicotinamide adenine dinucleotide (NAD<sup>+</sup>), L-malic acid, citrate synthase, malate dehydrogenase, and propionic anhydride were purchased from Sigma. Peptides were purchased from Peptide2.0. Zinc was purchased as an ICP standard from GFS Chemicals and the acetic acid standard was purchased from Ricca Chemical Company. Chelex 100 resin was purchased from Bio-Rad. Acetyl lysine was purchased from Chem-Impex Chemical International Inc. All other materials were purchased from Fisher and were of a purity >95 % unless otherwise noted.

## HDAC8 expression and purification

2xYT media supplemented with 50 µg/mL ampicillin was inoculated with *E. coli* cells (BL21 (DE3)pHD4-HDAC8-TEV-His<sub>6x</sub>) and grown overnight at 37°C. 12.5 mL overnight culture/L of expression culture was used to induce modified-autoinduction TB medium (12 g/L tryptone, 24 g/L yeast extract, 8.3g/L Tris-HCl, 4 g/L lactose, 1 g/L glucose, 10 mL/L glycerol, pH 7.4) supplemented with 50 µg/mL ampicillin and 200 µM ZnSO<sub>4</sub>. The cells were grown overnight at 30°C and harvested 20 - 24 hours post inoculation (9,000 x g, 10 min, 4°C). Cells were either lysed immediately or frozen as a dry cell pellet at -80°C. The cell pellet was resuspended in low salt DEAE buffer (50 mM HEPES, 200 µM ZnSO<sub>4</sub>, 1 mM TCEP, 50 mM NaCl, 5 mM KCl, 1 µg/mL tert-Amyl methyl ether (TAME), 10 µg/mL PMSF, pH 7.8) and passed 2 times through a microfluidizer (Microfluidics) to lyse. The lysate was then stirred on ice with 0.1% polyethylenimine (pH 7.9) for 10 - 15 minutes and pelleted (39,000 x g, 45 min, 4°C). The supernatant from 20 L of cell culture was loaded onto a 100 mL DEAE Sepharose column, and the column was washed with 300 mL low salt DEAE buffer. Protein was eluted with 200 mL high salt DEAE buffer (50 mM HEPES, 200 µM ZnSO<sub>4</sub>, 1 mM TCEP, 500 mM NaCl, 5 mM KCl, pH 7.8). The eluate was dialyzed against 50 mM HEPES, 50 mM NaCl, 25 mM imidazole, pH 7.8 for 1 hour at 4°C to decrease the salt and ZnSO<sub>4</sub>. 20 mL of chelating sepharose resin charged with NiSO<sub>4</sub> and equilibrated with imidazole buffer containing 50 mM HEPES, 100 mM NaCl, pH 7.8, 25 mM imidazole was incubated for 20 - 30 minutes stirring on ice with the dialyzed protein fraction. The slurry was poured into a column and allowed to drain. The column was then washed with a stepwise gradient using buffer supplemented with 50 mM, and 250 mM imidazole. The presence of HDAC8 was confirmed by SDS-PAGE analysis. These fractions were combined and dialyzed overnight in 25 mM imidazole buffer containing 1 mg His-Tagged TEV(S219V) protease per 15mg protein at 4°C overnight. The protein was applied to a chelating sepharose resin charged with NiSO<sub>4</sub>. The column was washed with 50mM imidazole buffer and the flow through was collected and concentrated to <3 mL. The protein was run over a GE HiPrep 16/60 Sephacryl S200 HR size exclusion column equilibrated with size exclusion buffer (30 mM HEPES, 150 mM NaCl, 1 mM TCEP, pH 8). The fractions of >95 % pure HDAC8 (confirmed by SDS-PAGE) were combined and concentrated to <0.7 mL. HDAC8 was dialyzed against 4 L of metal free chelation buffer (30 mM MOPS, 3 mM KCl, 1 mM EDTA, 1 mM

TCEP, pH 7.8) overnight at 4°C to remove metal ions from the proteins. Then, HDAC8 dialyzed against metal free dialysis buffer (30 mM MOPS, 3 mM KCl, 1 mM TCEP, pH 7.8) and incubated at 4°C overnight and run over a PD-10 column (GE) in metal free dialysis buffer. The fractions containing HDAC8 were concentrated to >500 μM, flash frozen in 10 μL aliquots, and stored at -80°C. HDAC8 activity was confirmed using the Fluor de Lys assay as described previously[32, 41].

### **HDAC assay**

Assays were performed as previously described[42] with a few modifications. Briefly, depending on their solubility, peptides were dissolved in water, 50% acetonitrile, or 10% DMSO. The dissolved peptide solutions were chelated by incubation with Chelex resin at 4°C for more than three hours. Peptide concentrations were measured by fluorescamine assay or OD<sub>280</sub> as previously described[42, 43]. Peptides (0-100 μM) were incubated in HDAC8 assay buffer (50 mM HEPES, 137 mM NaCl, 3.7 mM KCl, pH 7.8) for 10 minutes at 30°C before initiating the reactions with the addition of 0.5 μM Zn(II)-HDAC8. Assays were performed at 30°C. The reactions were worked up as previously described. Initial rates were fit to the linear portion of the time versus product curve.

### **Histone expression and purification**

H3 histones were expressed and purified as previously described[40] with a few modifications. Briefly, Quikchange PCR was utilized to modify the codon representing the Lys9, Lys14, and Lys56 sites in histone H3 (PCDF PyLT-1 plasmid) to an amber codon sequence. BL21 (DE3) cells were transformed with the mutant or wild type PCDF PyLT-1 and pAcKRS-3 plasmids for H3 expression (generous gifts from Jason Chin). To express H2A, H2B, and H4, expression plasmids of each histone were used (generous gifts from Geeta Narlikar).[44] BL21(DE3) cells transformed with the respective plasmids were grown in LB supplemented with the proper antibiotic (kanamycin and streptomycin for H3, or ampicillin for H2A, H2B, and H4) and were grown overnight at 37°C. Then the cultures were used to inoculate LB medium supplemented

with the proper antibiotic and incubated at 37°C until reaching an OD<sub>600</sub> of 0.7. For acetylated H3, 20 mM nicotinamide and 10 mM acetyl lysine was added, and 30 minutes later protein expression was induced by addition of 0.5 mM IPTG. For expression of the other histones, the cells were induced with 0.5 mM IPTG after reaching an OD<sub>600</sub> of 0.6. The cultures were harvested 3 - 4 hours after induction (9,000 x g, 10 min, 4°C). The cell pellets were stored at -80°C as a dry pellet.

Cell pellets were resuspended in Phosphate Buffered Saline (PBS) supplemented with 20 mM nicotinamide and lysed using the M110L Microfluidizer (Microfluidics). The lysate was centrifuged (39,000 x g, 45 min, 4°C) and the supernatant was discarded. The pellets were resuspended in PBS supplemented with Triton X-100 and 20 mM nicotinamide and centrifuged (39,000 x g, 45 min, 4°C). In total, the cell pellets were resuspended and centrifuged twice in PBS with Triton X-100 and nicotinamide, and twice in PBS with nicotinamide. The pellets were then macerated in DMSO and incubated at room temperature for 30 minutes to an hour. After the incubation, 6 M guanidinium chloroide, 20 mM Tris, and 2 mM dithiothreitol (DTT), pH 8 was added to the DMSO mix and incubated with shaking at 37°C for 1 hour. The rest of the protocol proceeded as previously described for H3 using two a nickel column, followed y cleavage of the His-tag with TEV protease and a second nickel column[40]. For H2A, H2B, and H4, the DMSO mix was centrifuged as above and dialyzed against 7 M urea, 100 mM sodium acetate (NaOAc), 20 mM Tris, 100 mM NaCl, pH 7.5 overnight at 4°C. The dialyzed protein extract was applied to an SP sepharose column, the column was washed (7 M urea, 100 mM NaOAc, pH 5.2) and protein was eluted stepwise with SP buffer (7 M urea, 100 mM NaOAc, pH 7.5) supplemented with 300 mM, 500 mM and 1 M NaCl. The fractions containing H2A, H2B, and H4 (confirmed by SDS-PAGE) were collected and dialyzed 2 times overnight against MilliQ water and 2 mM 2-mercaptoethanol (BME). After dialysis, the histones were lyophilized (Labconco Freezone Plus 6) and stored at -80°C. Tetramer, octamer, and nucleosome were reconstituted as previously described [40, 44].

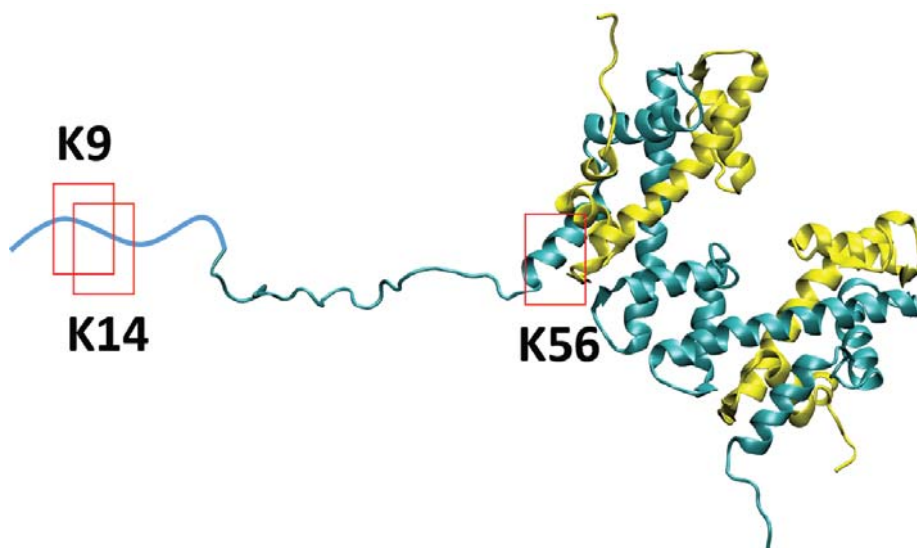
## Protein deacetylation assays

HDAC8 was reconstituted with 1:1 Zn(II) for 1 hour on ice in 1x HDAC8 assay buffer[32]. Histone complexes were incubated in HDAC8 assay buffer (above) for 10 minutes at 30°C before initiating reactions by addition of 0 - 15  $\mu$ M Zn(II)-HDAC8. Reactions were performed at 30°C. Reactions were quenched by addition of 25% trichloroacetic acid at varying times. The samples were incubated for 30 minutes on ice and then centrifuged (16,800 x g, 10 min, 4°C). The supernatant was removed and the pellet was washed 2 times with acetone and centrifuged (as above) after each wash. The pellet was dried in a SPD1010 SpeedVac System (ThermoSavant) and frozen at -80°C. The pellets were resuspended in 2  $\mu$ L propionic anhydride and ammonium hydroxide (NH<sub>4</sub>OH) and incubated at 51°C for 1 hour. 30  $\mu$ L of 50 mM NH<sub>4</sub>HCO<sub>3</sub> was added to each tube, and the pH of each tube was adjusted to 8 using NH<sub>4</sub>OH as indicated by pH paper. Then 0.2  $\mu$ g sequence grade trypsin (Promega) was added to each tube and incubated overnight at 37°C. 3.5  $\mu$ L of 10% formic acid was added to each tube, and the tubes were centrifuged (16,800 x g, 10 min, 4°C). The supernatant was removed and deacetylation was quantified by MS/MS mass spectrometry analysis in the lab of Andrew Andrews (Fox Chase Cancer Center) as previously described[45]. The kinetic data were analyzed using a single exponential decay:  $\frac{[\text{product}]}{[\text{substrate}]} = e^{-k_{\text{obs}} t}$ . Variations between days was less than ~20 %.

## Results

Numerous studies have sought to evaluate HDAC substrate specificity by measuring activity with acetylated peptides[10-13]. Utilizing these peptides to determine substrate specificity is predicated on the assumption that HDAC8 uses the same interactions to distinguish peptide substrates as it does for full length proteins. To test this assumption we probed three deacetylation sites on the putative HDAC8 substrate histone H3. Two biologically relevant sites (H3K9ac and H3K14ac) were chosen in proximity to each other on N-terminal histone tail as they share an unfolded secondary structure (Figure 5.1) but differ in sequence. These sites test the role of sequence in determining HDAC8 specificity. A third site (H3K56ac) was chosen

because it is contained within an  $\alpha$ -helix in the globular structure of H3 (Figure 5.1). This site allows probing of the role of secondary structure and long range protein-protein interactions with HDAC8 recognition. Varied concentrations of acetylated peptides containing 7 and 13 amino acids (7- and 13-mer peptides) representing the sequence of these three H3 acetylation sites were incubated with HDAC8 (Figure 5.2). The rate of deacetylation was measured using an assay that couples the formation of NADH to acetate[42]. Acetate product increased linearly up to 20 minutes. The initial rates for these reactions are linearly dependent on the peptide concentration yielding values for  $k_{cat}/K_M$ , as shown in Table 5.1.



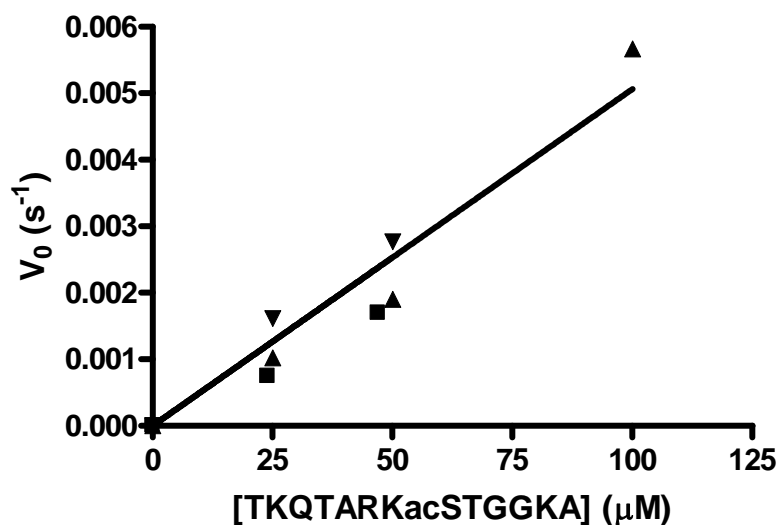
**Figure 5.1 Schematic of acetylation sites tested**

Shown is the crystal structure[44] of the H3 (blue) and H4 (yellow) histone dimer. The residues 1 to 20 are shown in an extended conformation as they have no discrete structure within the crystal structure. Three sites on the histone tetramer (a dimer of the H3/H4 dimers) were acetylated using non-natural amino acid incorporation into expressed proteins. The K9 and K14 sites are present in the histone tail while the K56 site is present in an  $\alpha$ -helical section within the globular section of the protein.



	H3 K9 site	H3 K14 site	H3 K56 site
7-mer	Ac-TARKacSTG-NH <sub>2</sub>	Ac-TGGKacAPR-NH <sub>2</sub>	Ac-RYQKacSTE-NH <sub>2</sub>
	56 ± 6.0	8.0 ± 0.69	78 ± 7.6
13-mer	Ac-TKQTARKacSTGGKA-NH <sub>2</sub>	Ac-RKSTGGKacAPRKQL-NH <sub>2</sub>	Ac-EIRRYQKacSTELLI-NH <sub>2</sub>
	51 ± 3.1	21 ± 3.5	100 ± 11
Tetramer	>45,000	2,500 ± 71	4,000 ± 570
Octamer		990 ± 200	

**Table 5.1 Catalytic efficiencies of substrates**



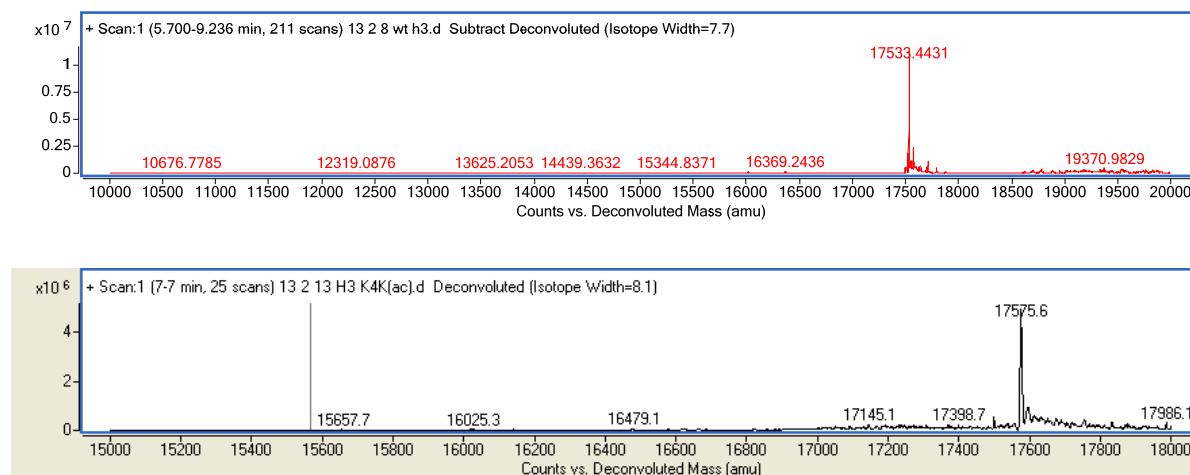
**Figure 5.2 Example substrate dependence for the 13-mer H3K9ac peptide**

Of the 7-mer peptides, the H3K56ac peptide demonstrated the highest HDAC8  $k_{cat}/K_M$  values ( $78 \text{ M}^{-1}\text{s}^{-1}$ ) followed by the H3K9ac ( $56 \text{ M}^{-1}\text{s}^{-1}$ ) and H3K14ac sites ( $8 \text{ M}^{-1}\text{s}^{-1}$ ). Few differences in  $k_{cat}/K_M$  values are observed between the 7- and 13-mer peptides (Table 1.1). The differences in catalytic efficiency between the 7- and 13-mer peptides for the H3K9ac and H3K56ac peptides is <25 % while the larger H3K14ac peptide increases activity <3-fold. A 17-mer peptide representing the H3K9ac site also showed less than a three-fold increase in  $k_{cat}/K_M$  ( $120 \pm 11 \text{ M}^{-1}\text{s}^{-1}$ ). These data suggest that the three amino acids in flanking both sides of the acetyl lysine largely dictate HDAC8 peptide preference at the H3K9ac site.

### Deacetylation of acetylated H3/H4 tetramers

To probe whether HDAC8-substrate interactions distal from the active site play a role in HDAC8 substrate recognition, we developed a method to assay full length protein substrates. One major challenge in identifying HDAC8 substrates is determining the rates of deacetylation for individual acetyl lysine sites, as HDAC substrates may have multiple acetyl lysines. To prepare proteins with a single acetyl lysine residue, we used the method developed by Jason Chin's group which incorporates non-natural amino acids into recombinantly expressed proteins[39, 40]. This method has previously been shown to yield large amounts of acetylated proteins

substrate through bacterial expression and can be easily scaled. The acetyl lysine is incorporated into expressed proteins at an amber codon site (TAG) using a tRNA-cognate tRNA synthetase pair. We incorporated amber codons at the K9, K14, and K56 sites of the histone H3 sequence in the PCDF PyLT-1 plasmid[39, 40]. Full length histone H3 with a single acetyl lysine residue were subsequently expressed and purified. Q-tof mass spectrometry of modified histone H3 demonstrated that >95% of purified protein contains a single acetyl lysine (Figure 5.3). To stabilize H3, stoichiometric H4 was added to form H3/H4 tetramers and the complex was purified by size exclusion chromatography.



**Figure 5.3 H3 Histones are singly acetylated**

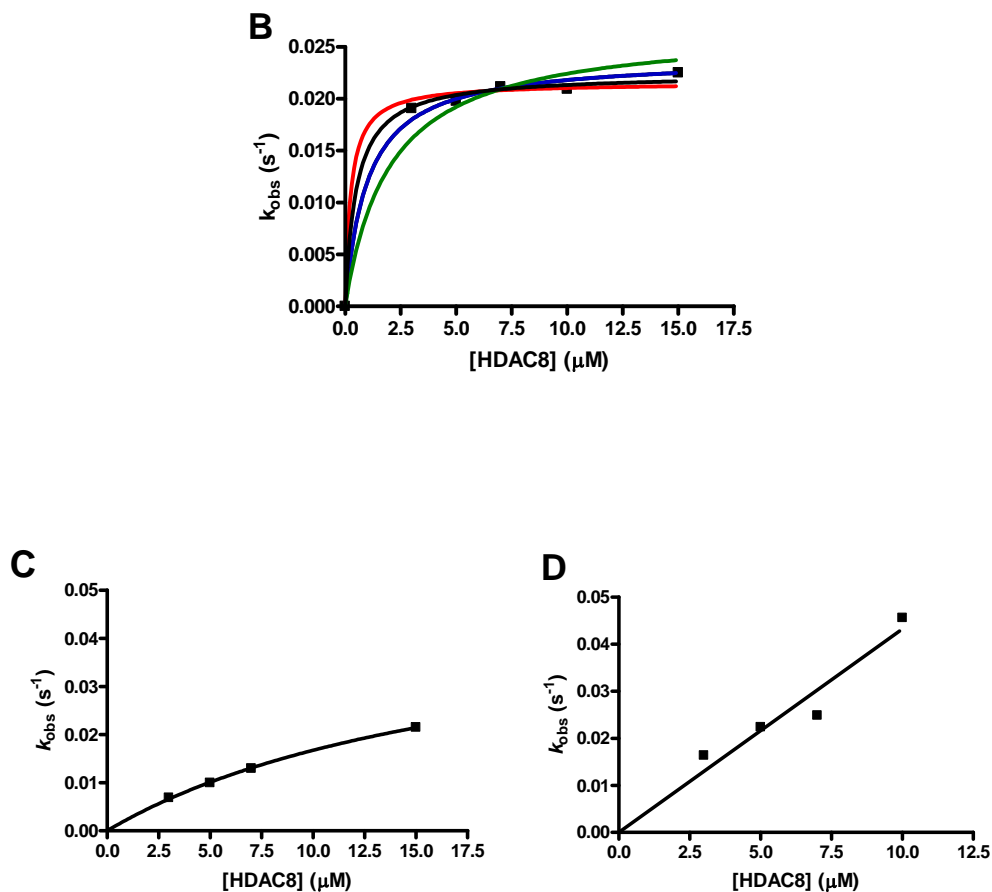
Q-Tof mass spectrometry traces showing unacetylated H3 (top) and singly acetylated H3 (bottom). These peaks are as expected based on the histone sequence and differ 42 m/z as would be expected for a singly acetylation. The location of the acetate moiety was confirmed using MS/MS (data not shown).

### Single turnover kinetics

To measure activity of HDAC8 with singly acetylated H3, we measured the kinetics under single turnover conditions ( $[HDAC8] \gg [H3]$ ) in which 0.5  $\mu$ M acetylated H3/H4 tetramer was incubated with 3 – 15  $\mu$ M HDAC8. The reactions were quenched with acid and the tetramer was trypsinized. The fraction of the acetylated peptide from H3 trypsinization was measured by

MS/MS mass spectrometry as a function of reaction time. These data were best described by an exponential equation (Figure 5.4a, b) yielding  $k_{obs}$  values. The  $k_{obs}$  values are linearly dependent on the HDAC8 concentration. Assuming rapid equilibration of the HDAC8-H3/H4 complex, a fit of the Michaelis-Menten equation to the data yielded  $k_{max}/K_{1/2}$  values equal to  $>45,000 \text{ M}^{-1}\text{s}^{-1}$ ,  $2500 \pm 71 \text{ M}^{-1}\text{s}^{-1}$ , and  $4000 \pm 570 \text{ M}^{-1}\text{s}^{-1}$  for the H3K9ac, H3K14ac, and H3K56ac tetramers, respectively. Each of these catalytic efficiencies is 10- to 100-fold faster than the corresponding peptide  $k_{cat}/K_M$  values, which suggests that interactions outside of the 13-mer peptide sequence are important for HDAC8 substrate recognition, assuming a comparable rate-limiting step (see discussion section). This increase in catalytic efficiency is mainly due to a decrease in  $K_{1/2}$  relative to  $K_M$ , suggesting enhanced binding of the protein substrates compared to the peptide. Because the deacetylation is likely the rate limiting step,  $K_{1/2}$  and  $K_M$  are approximately equal to  $K_D$ . While H3 peptides have  $K_M$  values higher than  $100 \mu\text{M}$ , (data not shown) the H3K9ac/H4 tetramer has a  $K_{1/2}$  value  $<0.5 \mu\text{M}$  and the H3K14ac/H4 tetramer has a  $K_{1/2} \sim 20 \mu\text{M}$ . These differences suggest that longer range interactions enhance reactivity of HDAC8 with full length substrate.

Finally, the 2-10 fold enhanced reactivity of HDAC8 with the H3K56ac peptide compared to the H3K9ac peptide is not maintained in the singly acetylated H3K9ac/H4 and H3K56ac/H4 tetramer, as would be expected if HDAC8 utilized sequence as the primary method of recognizing substrates. The ratio of reactivity of HDAC8 catalyzed deacetylation of acetylated protein vs acetylated peptide substrates varies from  $>800$  for the H3K9ac site to  $\sim 300$  for the H3K14ac site to  $\sim 50$  for the H3K56ac site. The increased reactivity for histone tail residues is consistent with enhanced accessibility of the H3K9ac/H4 compared to the H3K56ac/H4 site within the H3/H4 tetramer. In this case, the sequence of the H3K14ac/H4 site could also be a poor substrate site for deacetylation by HDAC8. Another explanation of decreased HDAC8 reactivity with the H3K56/H4 site is that the binding surface of an  $\alpha$ -helix is better represented by the  $\dots-3, 0, \text{ and } +3\dots$  residues and therefore a linear peptide may not represent these amino acids well.

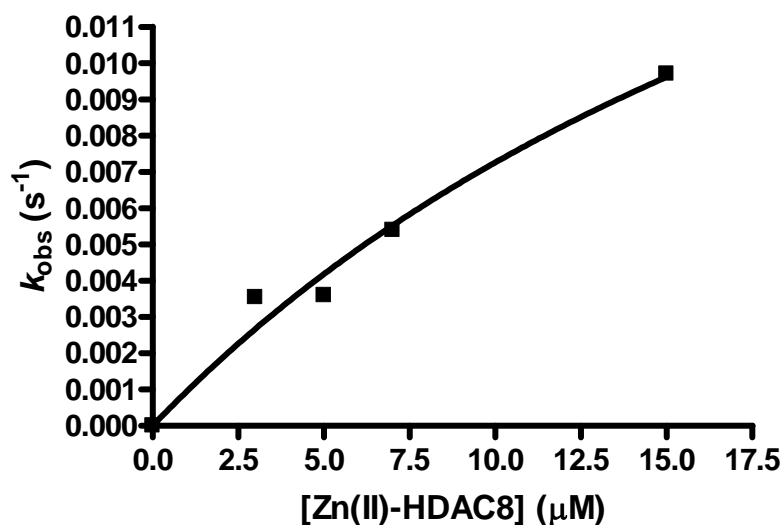


**Figure 5.4 Sample single turnover deacetylation data**

A) Sample data from a tetramer deacetylation with 7 μM HDAC8 and 0.5 μM H3K9ac/H4 tetramer (1 μM acetyl lysine). Data is best described by a single exponential fit. The residuals (below) from this single exponential fit show that a single exponential is within agreement of the data. B) H3K9ac/H4  $k_{obs}$  vs  $[HDAC8]$  curve. Three separate fits are present showing different potential  $K_{1/2}$  values:  $K_{1/2} = 0.25$  (red);  $K_{1/2} = 0.5$  (black);  $K_{1/2} = 1$  (blue);  $K_{1/2} = 2$  (green). Each fit has the respective  $k_{max}/K_{1/2}$  value: 86,000 M<sup>-1</sup>s<sup>-1</sup>; 45,000 M<sup>-1</sup>s<sup>-1</sup>; 24,000 M<sup>-1</sup>s<sup>-1</sup>; and 13,000 M<sup>-1</sup>s<sup>-1</sup> respectively. The fits below indicate that the  $K_{1/2}$  is less than 0.5 μM. C) H3K14ac/H4  $k_{obs}$  vs  $[HDAC8]$  curve. The exponential fit indicates that the  $k_{max}/K_{1/2}$  is 2,500 M<sup>-1</sup>s<sup>-1</sup>. D) H3K56ac/H4  $k_{obs}$  vs  $[HDAC8]$  curve. The exponential fit indicates that the  $k_{max}/K_{1/2}$  is 4,000 M<sup>-1</sup>s<sup>-1</sup>.

## Deacetylation of histone octamers

To determine whether the activity of HDAC8 is altered by the formation of larger histone complexes, we measured rates of deacetylation with octameric singly acetylated H3K14ac and compared this to HDAC8 deacetylation of tetramers. Histone octamers were reconstituted with two copies of each core histone (H2A, H2B, H3K14ac, and H4). The deacetylation rate catalyzed by HDAC8 was measured under single turnover conditions using mass spectrometry (as described above). The time-dependent decrease in the fraction of acetylated H3 was well described by a single exponential (Figure 5.4a). The resulting  $k_{\text{obs}}$  values vary with HDAC8 concentrations. A fit using the Michaelis-Menten equation yielded a  $k_{\text{max}}/K_{1/2}$  value of  $990 \pm 200 \text{ M}^{-1}\text{s}^{-1}$  for the H3K14ac octamer. This catalytic efficiency is decreased 3-fold compared to H3K14ac/H4 tetramer site and is ~45-fold faster than the deacetylation of H3K14ac peptides. These data suggests that the octamer is recognized very similarly to the tetramer for this site; the modest decrease in the  $k_{\text{max}}/K_{1/2}$  value likely reflects decreased accessibility for the acetylated lysine in the larger complex.

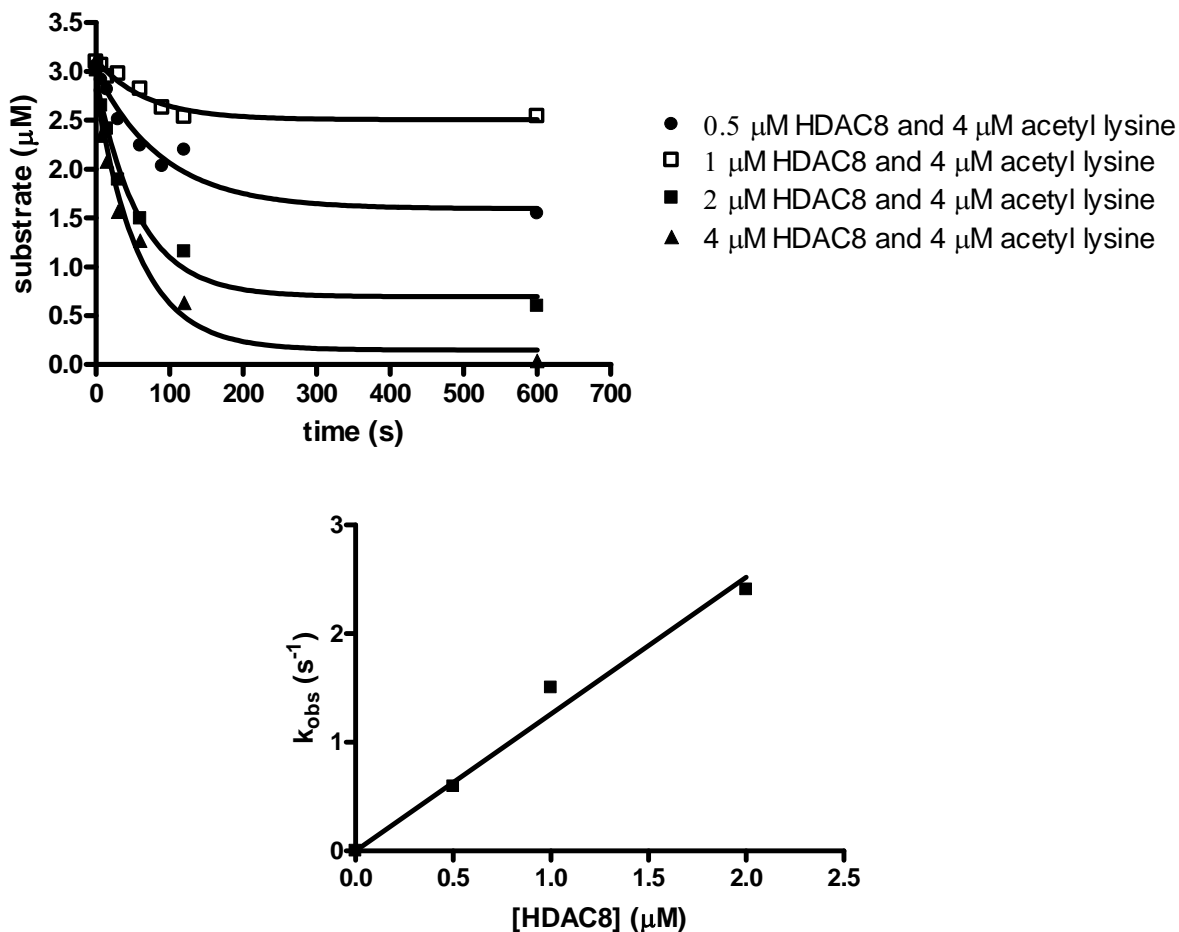


**Figure 5.5 H3K14ac octamer data**

H2a/H2b/H3K14ac/H4  $k_{obs}$  vs [HDAC8] curve. The fit indicates that the  $k_{max}/K_{1/2}$  is  $990 \text{ M}^{-1}\text{s}^{-1}$ .<sup>1</sup> This is within two-fold of the H3K14/H4 tetramer data.

### **Multiple turnover of tetramer**

Catalysis of deacetylation of a singly acetylated H3/H4 tetramers under multiple turnover conditions yielded unexpected results. Assays were performed with limiting HDAC8 and excess acetylated H3/H4 (0 - 2 μM HDAC8 and 4 μM acetyl lysine substrate). Deacetylation was assayed using the previously described mass spectrometric method. Unexpectedly, these reactions did not proceed to complete deacetylation of the acetylated H3/H4 tetramers as expected from the thermodynamics of the reaction. Incubation of 0.5 μM acetylated H3K9ac with 4 μM HDAC8 led to an initial decrease in acetylation that plateaued at 0.59 μM deacetylation product. Increasing the concentration of HDAC8 in the reaction led to both an increase in the initial rate of deacetylation and an increase in the plateau concentration of deacetylation product (Figure 5.6).



**Figure 5.6 HDAC8 turnovers vs enzyme concentration**

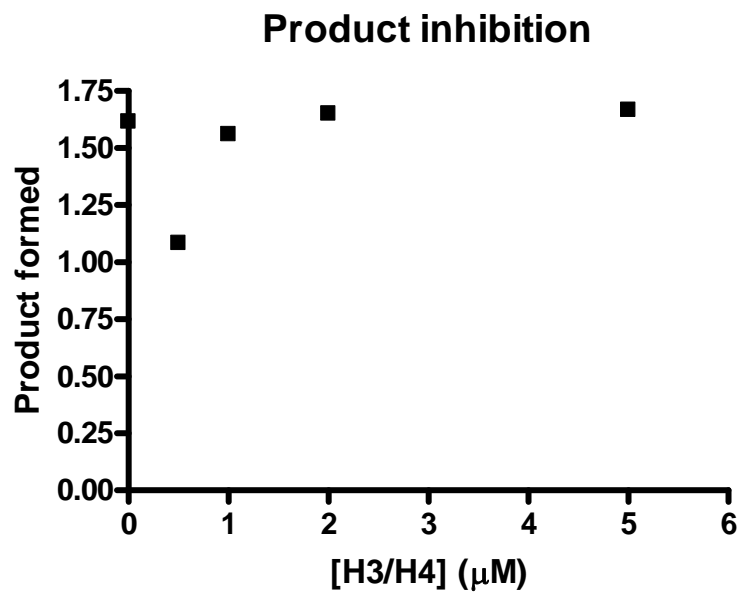
A) Raw data of the time versus substrate concentration. Substrate was kept constant while the concentration of HDAC8 was varied. Data was fit to an exponential curve. B) Plateaus from data in 4A graphed versus HDAC8 concentration. The slope of this curve indicates that  $1.2 \pm 0.069$  molecules of acetate are turned over per molecule of HDAC8. Fits were determined by a single reaction and errors were calculated based on the differences between the fit and the data.

The dependence of the initial rate on HDAC8 concentration is as expected for multiple turnover reactions ( $k_{obs} = \frac{k_{cat}[HDAC8][HACFad/H4]}{K_M + [HACFad/H4]}$ ). However the dependence of the apparent end point on the HDAC8 concentration is unexpected. A comparison of the end points to the HDCA8 concentration indicates that  $\sim 1$  mol product/mol HDAC8 was obtained in these reactions (Figure



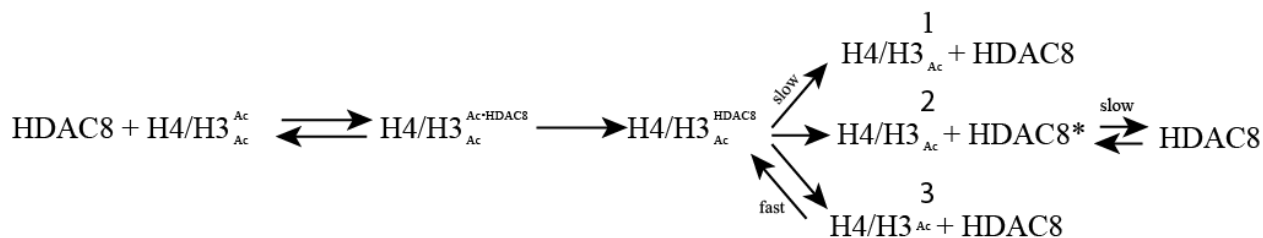
5.5b; slope =  $1.2 \pm 0.069$ ) before the reaction rate plateaued. One mechanism consistent with these data is that HDAC8 binds to acetylated H3/H4 substrate, catalyzes deacetylation and then product dissociation is very slow, leading to the observed biphasic kinetics. Alternatively, HDAC8 may dissociate after a deacetylation reaction but revert to a non-catalytically active form of the enzyme. Finally, the deacetylated H3/H4 tetramer may bind the HDAC8 with higher affinity than the substrate and inhibit further turnover.

To address the model of product inhibition, 0 – 5  $\mu\text{M}$  non-acetylated tetramer (product) was mixed with 2  $\mu\text{M}$  H3K9ac/H4 tetramer and 1  $\mu\text{M}$  HDAC8 was added to initiate the reaction. No inhibition of deacetylation was observed (Figure 5.7). Preincubation of HDAC8 and the product at 30°C or on ice for > 1 hour prior to the reaction also did not lead to inhibition of HDAC8 catalyzed deacetylation (results not shown). These data rule out the possibility that the H3/H4 product severely inhibits HDAC8. Therefore these data suggest that HDAC8 binds to substrate proteins, and catalyzes deacetylation. The reaction then stops at this point because HDAC8 is kinetically trapped to the product ceasing further deacetylations (Figure 5.8). Alternatively, HDAC8 may dissociate in a non-catalytically active enzyme form.



**Figure 5.7 No product inhibition measured**

4  $\mu\text{M}$  H3K9ac/H4, 1.5  $\mu\text{M}$  HDAC8, and 0 – 5  $\mu\text{M}$  H3/H4 (unacetylated) were incubated together at 30°C and the concentration of deacetylated tetramer was measured after 10 minutes of reaction. No product inhibition is measured for these reactions.



**Figure 5.8 Schematic of inhibition mechanism**

There are three potential inhibition mechanisms: 1) Dissociation from the HDAC·tetramer complex is slow; 2) Dissociation of the HDAC·tetramer complex is fast but HDAC8 comes off in a catalytically inactive form and isomerization into a catalytically active form is slow; 3) the deacetylated histone acts as an HDAC8 inhibitor. Data showing that unacetylated tetramer does not inhibit HDAC8 suggests that the 3<sup>rd</sup> mechanism is not correct.

## Discussion

Determining HDAC substrate recognition is important to understanding HDAC's impact on cellular regulation. Thus far HDAC recognition of protein substrates has been largely unexplored. For HDACs, the vast majority of studies have utilized small peptide substrates which typically span less than a 8 Å x 20 Å section of a ~45 Å x 45 Å binding surface[35]. This leaves a large number of potential protein recognition hotspots and negative interactions sites unexplored. The potential for many more interactions between HDAC8 and substrates exists within this larger binding surface. With a peptide, a single interaction of 0.5 - 2 kcal/mol can alter the catalytic efficiency 50-fold[42]. With a larger substrate, there are many more potential interaction sites, which can easily overcome the 2 kcal/mol of energy from local interactions. There is also precedence in the literature for HDAC8 utilization of interactions distal of the active site. A KRHR motif (present in the H4 histone) upstream of the acetyl lysine site, has been shown to increase HDAC8 catalyzed deacetylation rates of peptide substrates[12]. We therefore probed activity with acetylated protein substrates to determine whether longer range interactions form in additional recognition motifs for HDAC8.

Optimally, the  $k_{cat}/K_M$  value of the proteins would be compared directly to the  $k_{cat}/K_M$  values of the peptides. Due to burst kinetics measured for the multiple turnover of protein substrates, we instead measured single turnover kinetics. Comparisons between the single turnover kinetics ( $k_{max}/K_{1/2}$ ) of the H3/H4 tetramer and the multiple turnover of the peptides is possible due to the likelihood that product release is not rate limiting for the peptide substrates. Experimental results show that trifluoroacetate peptide substrates are deacetylated with faster chemistry ( $k_{cat}$ ) paralleling the higher chemical reactivity and suggests that product release is not rate limiting[28].  $k_{cat}/K_M$  values significantly slower than diffusion control ( $10^5 \text{ M}^{-1}\text{s}^{-1}$  vs  $10^0\text{-}10^5 \text{ M}^{-1}\text{s}^{-1}$ ) suggest that binding is the rate limiting step, and is further consistent with product release as the non-rate limiting step. The  $k_{max}/K_{1/2}$  value of the protein substrates can be compared to the  $k_{cat}/K_M$  values of the peptides.

Initial experiments were performed using peptides substrates for comparison with the acetylated full length acetylated proteins. Previous studies using peptides have shown that HDAC8 has a preference for aromatic amino acids in the position on the C-terminal side of the acetyl lysine (+1 position)[10, 11]. Based on this metric alone, the peptides representing the H3 histone are not predicted to be efficient HDAC8 substrates. As expected, these peptides were experimentally not efficient substrates with catalytic efficiencies ranging from poor ( $10^0 \text{ M}^{-1}\text{s}^{-1}$ ) to mediocre ( $10^2 \text{ M}^{-1}\text{s}^{-1}$ ) (Table 2.1). The interactions between 7-mer peptides occur within a  $\sim 10\text{\AA}$  radius of the active site and resulted in  $\sim 0.5 \text{ kcal/mol}$  (assuming the entropy change is minimal)<sup>1</sup>. The acetylated H3/H4 tetramers, which also lack an aromatic amino acid in the +1 position, were deacetylated 2 - 45 times more rapidly than the corresponding peptides. These results suggests that interactions present in the tetramer but absent from the peptides enhance HDAC8 recognition of protein substrates. The catalytic efficiency of the acetylated H3/H4 tetramer was increased by 3-6 kcal/mol, lowering the activation energy<sup>1</sup> suggesting that more significant and longer range interactions are responsible for HDAC8's recognition of proteins. The surprisingly large  $k_{max}/K_{1/2}$  values for the H3K9ac/H4 tetramer may be due to either a fairly strong interaction or an abundance of weak interactions. Additionally, the lowered reactivity of the H3K56ac substrates compared to the local and distal sequences as has been previously described[12] or may reflect a preference for acetyl lysines close to the C-terminus and/or enhanced accessibility.

---

<sup>1</sup>  $\Delta \Delta H \sim \Delta \Delta G = RT \cdot \ln((k_{cat}/k_{M1}) / (k_{cat}/k_{M2}))$

This second hypothesis is bolstered due to the lack of major differences observed for peptide substrates as they are all close to the termini of the substrate. Furthermore, histone octamers were deacetylated with similar catalytic efficiencies to the tetrameric substrates suggesting that interactions with the tetramer are sufficient to explain HDAC8 substrate interactions.

In many crystal structures, HDAC8 dimerizes at the substrate binding interface as part of the fundamental crystal unit. These structures may provide information about HDAC8 substrate recognition from the interactions between the two HDAC8 proteins. The HDAC8 substrate binding interface is a flexible 45 Å x 45 Å surface that has the potential to foster dozens of interactions. Crystal structures show van der Waals and hydrogen bonds between the HDAC8 dimers. The 3-6 kcal/mol difference measured for the peptides and tetramer associations could be due to 10 Van der Waals interactions or 6 hydrogen bonds, which are present in the diametric crystal structures. Also present among the interactions observed in the dimeric HDAC8 structure are repulsive interactions such as repulsive charge-charge interactions. As a result, attractive and repulsive interactions may work in concert to determine the HDAC8 substrate specificity. These additional interactions observed between these two HDAC8 subunits may explain the differences observed between the peptide and full length protein catalytic efficiencies.

Multiple turnover experiments with histone tetramer, showed that HDAC8 activity decreased after one cycle of deacetylation. These data best suggest the formation of a slowly reacting intermediate after the formation of the E·H3/H4·acetate complex. As HDAC8 is not inhibited by unacetylated tetramer, it suggests that HDAC8 is regulated by one of two mechanisms (Figure 5.8): Either dissociation of the products from the HDAC8 product complex is slow or HDAC8 forms an alternative inactive conformation after product formation. Further investigation of this phenomenon is necessary to determine whether this single turnover is a regulatory mechanism for HDACs *in vivo*, as this may prove to be a novel regulatory mechanism for HDAC8.

In this chapter I showed that HDAC8 catalyzes deacetylation of protein substrates at a rate constant >10-fold faster than peptide substrates. This difference in catalytic efficiency likely represents HDAC8-protein interactions which are absent in HDAC8-peptide binding, and likely affects binding. Furthermore, we have shown that slow product release or conformational change in HDAC8 may regulate multiple turnover reactions. These phenomena need further investigation and set the precedence for establishing initial parameters for investigation of

HDAC8 protein substrate specificity. In the future, different substrates and/or other HDAC complexes will be necessary to further define HDAC substrate specificity.

## **Bibliography**

- [1] S. Khochbin, A. Verdel, C. Lemerrier, D. Seigneurin-Berny, Functional significance of histone deacetylase diversity, *Curr Opin Genet Dev*, 11 (2001) 162-166.
- [2] X.J. Yang, E. Seto, The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men, *Nat Rev Mol Cell Biol*, 9 (2008) 206-218.
- [3] M.A. Glozak, N. Sengupta, X. Zhang, E. Seto, Acetylation and deacetylation of non-histone proteins, *Gene*, 363 (2005) 15-23.
- [4] C. Choudhary, C. Kumar, F. Gnad, M.L. Nielsen, M. Rehman, T.C. Walther, J.V. Olsen, M. Mann, Lysine acetylation targets protein complexes and co-regulates major cellular functions, *Science*, 325 (2009) 834-840.
- [5] S. Zhao, W. Xu, W. Jiang, W. Yu, Y. Lin, T. Zhang, J. Yao, L. Zhou, Y. Zeng, H. Li, Y. Li, J. Shi, W. An, S.M. Hancock, F. He, L. Qin, J. Chin, P. Yang, X. Chen, Q. Lei, Y. Xiong, K.L. Guan, Regulation of cellular metabolism by protein lysine acetylation, *Science*, 327 (2010) 1000-1004.
- [6] S.R. D'Mello, Histone deacetylases as targets for the treatment of human neurodegenerative diseases, *Drug news & perspectives*, 22 (2009) 513-524.
- [7] G. Li, H. Jiang, M. Chang, H. Xie, L. Hu, HDAC6 alpha-tubulin deacetylase: a potential therapeutic target in neurodegenerative diseases, *Journal of the neurological sciences*, 304 (2011) 1-8.
- [8] M.A. Glozak, E. Seto, Histone deacetylases and cancer, *Oncogene*, 26 (2007) 5420-5432.
- [9] P. Marks, R.A. Rifkind, V.M. Richon, R. Breslow, T. Miller, W.K. Kelly, Histone deacetylases and cancer: causes and therapies, *Nature reviews. Cancer*, 1 (2001) 194-202.
- [10] Z.A. Gurard-Levin, K.A. Kilian, J. Kim, K. Bahr, M. Mrksich, Peptide arrays identify isoform-selective substrates for profiling endogenous lysine deacetylase activity, *ACS Chem Biol*, 5 (2010) 863-873.

- [11] Z.A. Gurard-Levin, J. Kim, M. Mrksich, Combining mass spectrometry and peptide arrays to profile the specificities of histone deacetylases, *Chembiochem*, 10 (2009) 2159-2161.
- [12] Z.A. Gurard-Levin, M. Mrksich, The activity of HDAC8 depends on local and distal sequences of its peptide substrates, *Biochemistry*, 47 (2008) 6242-6250.
- [13] D. Riester, C. Hildmann, S. Grunewald, T. Beckers, A. Schwienhorst, Factors affecting the substrate specificity of histone deacetylases, *Biochem Biophys Res Commun*, 357 (2007) 439-445.
- [14] B.J. Wilson, A.M. Tremblay, G. Deblois, G. Sylvain-Drolet, V. Giguere, An acetylation switch modulates the transcriptional activity of estrogen-related receptor alpha, *Mol Endocrinol*, 24 (2010) 1349-1358.
- [15] M.A. Deardorff, M. Bando, R. Nakato, E. Watrin, T. Itoh, M. Minamino, K. Saitoh, M. Komata, Y. Katou, D. Clark, K.E. Cole, E. De Baere, C. Decroos, N. Di Donato, S. Ernst, L.J. Francey, Y. Gyftodimou, K. Hirashima, M. Hullings, Y. Ishikawa, C. Jaulin, M. Kaur, T. Kiyono, P.M. Lombardi, L. Magnaghi-Jaulin, G.R. Mortier, N. Nozaki, M.B. Petersen, H. Seimiya, V.M. Siu, Y. Suzuki, K. Takagaki, J.J. Wilde, P.J. Willems, C. Prigent, G. Gillessen-Kaesbach, D.W. Christianson, F.J. Kaiser, L.G. Jackson, T. Hirota, I.D. Krantz, K. Shirahige, HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle, *Nature*, 489 (2012) 313-317.
- [16] K.L. Durst, B. Lutterbach, T. Kummalue, A.D. Friedman, S.W. Hiebert, The inv(16) fusion protein associates with corepressors via a smooth muscle myosin heavy-chain domain, *Mol Cell Biol*, 23 (2003) 607-619.
- [17] J. Gao, B. Siddoway, Q. Huang, H. Xia, Inactivation of CREB mediated gene transcription by HDAC8 bound protein phosphatase, *Biochem Biophys Res Commun*, 379 (2009) 1-5.
- [18] P. Joshi, T.M. Greco, A.J. Guise, Y. Luo, F. Yu, A.I. Nesvizhskii, I.M. Cristea, The functional interactome landscape of the human histone deacetylase family, *Molecular systems biology*, 9 (2013) 672.
- [19] D. Waltregny, L. De Leval, W. Glenisson, S. Ly Tran, B.J. North, A. Bellahcene, U. Weidle, E. Verdin, V. Castronovo, 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*, 165 (2004) 553-564.
- [20] D. Waltregny, W. Glenisson, S.L. Tran, B.J. North, E. Verdin, A. Colige, V. Castronovo, Histone deacetylase HDAC8 associates with smooth muscle alpha-actin and is essential for smooth muscle cell contractility, *FASEB J*, 19 (2005) 966-968.

- [21] H. Lee, N. Sengupta, A. Villagra, N. Rezai-Zadeh, E. Seto, Histone deacetylase 8 safeguards the human ever-shorter telomeres 1B (hEST1B) protein from ubiquitin-mediated degradation, *Mol Cell Biol*, 26 (2006) 5259-5269.
- [22] A.J. de Ruijter, A.H. van Gennip, H.N. Caron, S. Kemp, A.B. van Kuilenburg, Histone deacetylases (HDACs): characterization of the classical HDAC family, *Biochem J*, 370 (2003) 737-749.
- [23] E. Hu, Z. Chen, T. Fredrickson, Y. Zhu, R. Kirkpatrick, G.F. Zhang, K. Johanson, C.M. Sung, R. Liu, J. Winkler, Cloning and characterization of a novel human class I histone deacetylase that functions as a transcription repressor, *J Biol Chem*, 275 (2000) 15254-15264.
- [24] J.J. Buggy, M.L. Sideris, P. Mak, D.D. Lorimer, B. McIntosh, J.M. Clark, Cloning and characterization of a novel human histone deacetylase, HDAC8, *Biochem J*, 350 Pt 1 (2000) 199-205.
- [25] I. Van den Wyngaert, W. de Vries, A. Kremer, J. Neefs, P. Verhasselt, W.H. Luyten, S.U. Kass, Cloning and characterization of human histone deacetylase 8, *FEBS Lett*, 478 (2000) 77-83.
- [26] A. Saha, G.N. Pandian, S. Sato, J. Taniguchi, K. Hashiya, T. Bando, H. Sugiyama, Synthesis and biological evaluation of a targeted DNA-binding transcriptional activator with HDAC8 inhibitory activity, *Bioorg Med Chem*, 21 (2013) 4201-4209.
- [27] H. Lee, N. Rezai-Zadeh, E. Seto, Negative regulation of histone deacetylase 8 activity by cyclic AMP-dependent protein kinase A, *Mol Cell Biol*, 24 (2004) 765-773.
- [28] B.C. Smith, J.M. Denu, Acetyl-lysine analog peptides as mechanistic probes of protein deacetylases, *J Biol Chem*, 282 (2007) 37256-37265.
- [29] E. Verdin, *Histone deacetylases : transcriptional regulation and other cellular functions*, Humana Press, Totowa, N.J., 2006.
- [30] D.P. Dowling, S.L. Gantt, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structural studies of human histone deacetylase 8 and its site-specific variants complexed with substrate and inhibitors, *Biochemistry*, 47 (2008) 13554-13563.
- [31] D.P. Dowling, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structures of metal-substituted human histone deacetylase 8 provide mechanistic inferences on biological function, *Biochemistry*, 49 (2010) 5048-5056.



- [32] S.L. Gantt, S.G. Gattis, C.A. Fierke, Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion, *Biochemistry*, 45 (2006) 6170-6178.
- [33] S.L. Gantt, C.G. Joseph, C.A. Fierke, Activation and inhibition of histone deacetylase 8 by monovalent cations, *J Biol Chem*, 285 (2010) 6036-6043.
- [34] A. Vannini, C. Volpari, G. Filocamo, E.C. Casavola, M. Brunetti, D. Renzoni, P. Chakravarty, C. Paolini, R. De Francesco, P. Gallinari, C. Steinkuhler, S. Di Marco, Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor, *Proc Natl Acad Sci U S A*, 101 (2004) 15064-15069.
- [35] A. Vannini, C. Volpari, P. Gallinari, P. Jones, M. Mattu, A. Carfi, R. De Francesco, C. Steinkuhler, S. Di Marco, Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex, *EMBO Rep*, 8 (2007) 879-884.
- [36] J.R. Somoza, R.J. Skene, B.A. Katz, C. Mol, J.D. Ho, A.J. Jennings, C. Luong, A. Arvai, J.J. Buggy, E. Chi, J. Tang, B.C. Sang, E. Verner, R. Wynands, E.M. Leahy, D.R. Dougan, G. Snell, M. Navre, M.W. Knuth, R.V. Swanson, D.E. McRee, L.W. Tari, Structural snapshots of human HDAC8 provide insights into the class I histone deacetylases, *Structure*, 12 (2004) 1325-1334.
- [37] M. Marek, S. Kannan, A.T. Hauser, M. Moraes Mourao, S. Caby, V. Cura, D.A. Stolfa, K. Schmidtkunz, J. Lancelot, L. Andrade, J.P. Renaud, G. Oliveira, W. Sippl, M. Jung, J. Cavarelli, R.J. Pierce, C. Romier, Structural basis for the inhibition of histone deacetylase 8 (HDAC8), a key epigenetic player in the blood fluke *Schistosoma mansoni*, *PLoS Pathog*, 9 (2013) e1003645.
- [38] L. Whitehead, M.R. Dobler, B. Radetich, Y. Zhu, P.W. Atadja, T. Claiborne, J.E. Grob, A. McRiner, M.R. Pancost, A. Patnaik, W. Shao, M. Shultz, R. Tichkule, R.A. Tommasi, B. Vash, P. Wang, T. Stams, Human HDAC isoform selectivity achieved via exploitation of the acetate release channel with structurally unique small molecule inhibitors, *Bioorg Med Chem*, 19 (2011) 4626-4634.
- [39] H. Neumann, S.Y. Peak-Chew, J.W. Chin, Genetically encoding N(epsilon)-acetyllysine in recombinant proteins, *Nature chemical biology*, 4 (2008) 232-234.
- [40] H. Neumann, S.M. Hancock, R. Buning, A. Routh, L. Chapman, J. Somers, T. Owen-Hughes, J. van Noort, D. Rhodes, J.W. Chin, A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation, *Molecular cell*, 36 (2009) 153-163.
- [41] D. Wegener, F. Wirsching, D. Riester, A. Schwienhorst, A fluorogenic histone deacetylase assay well suited for high-throughput activity screening, *Chemistry & biology*, 10 (2003) 61-68.

[42] N.A. Wolfson, C.A. Pitcairn, E.D. Sullivan, C.G. Joseph, C.A. Fierke, An enzyme-coupled assay measuring acetate production for profiling histone deacetylase specificity, *Anal Biochem*, (2014).

[43] X. Huang, M. Hernick, A fluorescence-based assay for measuring N-acetyl-1-D-myo-inositol-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase activity, *Anal Biochem*, 414 (2011) 278-281.

[44] K. Luger, A.W. Mader, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature*, 389 (1997) 251-260.

[45] Y.M. Kuo, A.J. Andrews, Quantitating the specificity and selectivity of Gcn5-mediated acetylation of histone H3, *PLoS One*, 8 (2013) e54896.

## **Chapter 6**

### **Conclusions**

To date, work determining HDAC substrate specificity has been predicated on a number of assumptions including: 1) Full length protein substrates are deacetylated with the same specificity as peptides sharing the same sequence. 2) The two amino acids flanking the acetyl lysine are important to HDAC specificity. 3) Methylcoumarin moieties do not alter peptide specificity. These assumptions have allowed researchers to simplify complex HDAC substrate specificity studies. Yet, these studies have failed to yield significant insights into biologically relevant substrates. As a result, the focus of my thesis has been to evaluate the validity of these assumptions. In my final chapter, I theorize how I would continue to move forward addressing the protein substrates and biological function of HDACs.

#### **Differentiating substrates from binding partners and downstream effectors**

One major issue plaguing the HDAC field is determining whether a protein associated with HDAC or affected by perturbation of HDAC activity is a substrate, binding partner, or downstream target. Assays for parsing HDAC interactions have utilized a series of techniques such as association studies (2-hybrid[1] and pull downs[1-3]), functional assays (knockdowns and overexpression[4-6]), and motif identification studies[7-9]. While these methods have produced an array of proteins (Chapter 1) they have yet to identify (on a large scale) the relationship between HDAC and these proteins. As a result, there are likely many misidentified

substrates and binding partners of HDAC8. In my research, (Chapter 4) it is unclear whether the acetylated proteins identified by mass spectrometry after inhibition of HDAC8 are direct substrates of HDAC8 or downstream targets affected by HDAC8 inhibition. As a result, these proteins cannot be positively identified as substrates.

In another example from my thesis, I have described data demonstrating that HDAC8 catalyzes deacetylation of H3 histones which suggests that H3 may be an *in vivo* substrate of HDAC8 (Chapter 5). This result needs to be validated by *in vivo*. Experiments show that H3 acetylation levels differ modestly upon HDAC8 overexpression (add ref). These results can be juxtaposed to results showing that H4 acetylation levels are significantly perturbed upon overexpression of HDAC8[5]. These data suggest that HDAC8 reacts with H4 more efficiently than H3 *in vivo*. However there are no data directly confirming this hypothesis. An alternative explanation, proposed in Chapter 5, is that HDAC8 may deacetylate both H4 and H3 but that HDAC8 deacetylates H3 only in specific gene regions, which mutes the observation of acetylation changes upon varying the HDAC8 concentration. Another alternative is that acetylation levels of both proteins are regulated by other HDAC8 substrates. One way to begin to answer this question is to prepare singly acetylated H4 histones and measure the reactivity of HDAC8 with the acetylated H3/H4 tetramer using the methods described in Chapter 5. We predict that HDAC8 would deacetylate H4 at least an order of magnitude faster than H3, if it is a better *in vivo* substrate. Caveats with this methodology are that the *in vitro* deacetylation rate may not be the only factor affecting HDAC8 activity *in vivo* and that deacetylation activity is affected by multiple post-translational modifications on the histone tails (see Chapter 1).

There are a large number of proposals as to how HDAC8 selects its substrates, which include cofactors that alter molecular recognition and regulation of multiple turnover activity by cellular release factors (Chapter 5). These hypotheses suggest that HDAC8 interacts with and/or affects a number of non-substrate proteins. In all, many careful experiments will ultimately be needed to identify the HDAC8 biological substrates. These experiments likely include those performed in my thesis. Other experiments such as the bump-and-hole [10] method or crosslinking methods[11-13] may also be helpful in identifying HDAC8 substrates and differentiating them from binding partners and proteins affected downstream by HDAC perturbation. In the

meantime, we must be cautious when interpreting our results as changes in phenotypes due to HDAC8 alterations may occur as a result of indirect HDAC effects on those proteins.

### **Identifying the local HDAC-substrate contacts using peptides**

Another issue affecting the HDAC field has been the use of peptide substrates containing a variety of N- and C- terminal moieties. These peptides have been used to make conclusions about substrate specificity, though it is unclear how accurately these peptides report on the recognition of full length proteins[7-9, 14, 15]. To date, these studies have been largely ineffective in determining the substrate sets of HDACs.

The Fluor de Lys assay[16] is one of the main activity assays used by the HDAC field, yet its ability to accurately profile substrate selectivity remains unclear. This assay utilizes a peptide conjugated to a methylcoumarin moiety on the C-terminal side of the acetyl lysine residue (Figure 1.2a). Upon deacetylation, this peptide becomes a trypsin substrate which allows the methylcoumarin to be cleaved from the peptide. The cleaved and uncleaved methylcoumarin moieties fluoresce at different wavelengths which allows deacetylation to be quantified.

Kinetic and substrate specificity studies on peptides containing a methylcoumarin have been largely predicated on the assumption that the methylcoumarin moiety does not alter substrate selectivity. Studies by the Mrksich lab suggest that this is not the case[8]. These studies show that HDAC8 deacetylates a methylcoumarin-conjugated peptide >10-fold faster than the same peptide lacking the methylcoumarin[8], and it is unclear whether this alters selectivity. As the SAMDI mass spectrometry method used to measure HDAC8 reactivity is semi-quantitative, I developed a coupled assay to measure HDAC8 catalyzed deacetylation of unlabeled peptides (Chapter 2). Using the coupled assay, I was able to show that the observations made by the Mrksich group were correct, though the quantitative changes were previously overestimated. Only a ~2-fold change was observed in catalytic efficiencies between peptides with a methylcoumarin and amide C-terminus, although larger differences were observed between other moieties (Chapter 2). For instance, substitution of a carboxylate moiety for the methylcoumarin at the C-terminus decreases the value of  $k_{cat}/K_M$  by >40-fold. This >40-fold change corresponds

to ~2.2 kcal/mol and shows that a few interactions between HDAC8 and a peptide substrate can drastically alter substrate specificity. While these results could be peptide specific, they suggest that HDAC8 is very sensitive to changes in substrate sequence. However, these C-terminal elements do not exist in a full length protein and likely do not play an important role in HDAC8 selectivity *in vivo*. Therefore standardization of the peptide termini is important for using peptides to dissect HDAC substrate specificity. Interestingly, peptide length (7 to 15 amino acids) barely altered substrate selectivity (Chapter 5). Taken together these results suggest that HDAC8 recognizes specific positions in the peptide sequence, particularly the 2 amino acids flanking both sides of the acetyl-lysine moiety. This phenomenon has been previously alluded to but has not yet swayed the field to standardize substrate length and termini[8, 15].

Finally, data performed in collaboration with Ed Holson's lab (Chapter 4) suggests that HDAC8 protein substrates may yield sequences that are good peptide substrates. The Holson lab utilized a unique mass spectrometry approach to identify possible *in vivo* HDAC8 protein substrates. Peptides representing the protein substrates identified by the mass spectrometry method were assayed. Two of the peptides identified by this method (ARAD1A and CSRP2BP) represent two of the fastest known HDAC8 peptide substrates. Further studies using computational methods to identify efficient peptide substrates in collaboration with Ora Schueler Furman's lab (Chapter 3), will be useful in determining whether the converse is true: that peptides can be used to predict HDAC *in vivo* substrates. Currently, we are in the process of using the algorithm from the Schueler-Furman lab, in conjunction with Ed Holson's mass spectrometry approach, to identify *in vivo* HDAC8 substrates. Validation that efficient peptide substrates identify *in vivo* substrates would be an important step forward for the HDAC field, indicating that these will be an effective means to determine HDAC substrate sets. Mass spectrometric experiments performed by Ed Holson's group (Chapter 4) are not sufficiently sensitive to detect the acetylated proteins that the Schueler-Furman algorithm predicts as HDAC8 substrates (Chapter 3). We are therefore developing techniques to isolate and test targets identified by the computational approach such as protein specific pull downs.

## Utilizing proteins to identify long range HDAC-substrate interactions

The putative substrate binding surface of HDAC8 extends tens of angstroms beyond the peptide binding site and thus it is reasonable to propose that longer range interactions are involved in substrate binding/specificity. To date there are no studies which have determined detailed kinetics on the deacetylation of full length proteins. In my thesis I present the first detailed kinetic study of HDAC8-catalyzed deacetylation of protein substrates (Chapter 5) which is an important step in determining *in vivo* HDAC8 substrate preferences. The deacetylation rate constants ( $k_{cat}/K_M$ ) catalyzed by HDAC8 for a protein substrate is 1,000-fold larger than the corresponding peptide substrate. As a result, these studies strongly suggest that HDAC8 utilizes distal interactions to enhance substrate affinity. These data suggests that a new set of long range interactions will need to be discovered to completely understand HDAC8 substrate specificity. Within a single protein (the H3 histone) kinetic measurements show that location in addition to sequence can significantly alter substrate recognition. In the H3 histone for instance, deacetylation of the H3K9ac/H4 site as opposed to the H3K14ac/H4 site (only 5 amino acids C-terminal of the H3K9ac/H4 site) shows a 10-fold change in catalytic efficiency. This is in contrast to peptides which show that longer sequences (7- vs 17-mer) do not significantly alter substrate specificity. Further experiments using other HDAC substrates (e.g., SMC3, CREB, and ARID1A) will be necessary to determine whether the selectivity rates identified for the H3 histone extend to other protein substrates. Also understanding how sequence and position play a role will be important in understanding how HDACs recognize their substrates. As a result switching replacing the sequences in the histone tails with the best computationally determined peptides will further reveal how short range and long range interactions dictate HDAC specificity.

Work within my thesis suggests that for the H3/H4 histone tetramer, HDAC8 binds to the globular surface of H3 and then utilizes short range interactions to define whether a specific acetylated lysine is a substrate (Chapter 5). It is interesting to consider that in the case of the mass spectrometry approach only a limited number of the identified acetylated proteins correspond to good peptide substrates (Chapter 4). One explanation for these data may simply be that acetylation of these proteins may be regulated by downstream effects of HDAC8 activity. Alternatively, long range interactions and/or protein cofactors may define substrate specificity

for these substrates making the peptides poor substrates. In this case a full length substrate would need to be assayed to show that these are *in vivo* substrates. In the future, assaying the full length proteins identified in Chapter 4 may reveal whether this is the case.

The results presented in Chapter 5 suggest that evaluation of contacts distal to the peptide binding site will be key to determining the *in vivo* substrate specificity of HDAC8. It is currently unclear whether these results obtained for the reactivity of HDAC8 with protein substrates will extend to other HDAC isozymes, and it will be interesting to discover whether each HDAC utilizes a similar or a unique set of interactions to recognize their substrates. Further analysis of HDAC8 with defined acetylated proteins is essential for delineating the molecular recognition of these proteins.

Finally, full length proteins have been integral in revealing a potential novel regulatory mechanism for HDAC8. Upon deacetylation of a full length histone, HDAC8 only performed one cycle of deacetylation before stopping. Further experiments on other protein substrates will be necessary to determine whether this mechanism is utilized *in vivo*. Should it be used, HDAC8 regulation may depend on small molecules or complexes as another means of regulation.

### **Does HDAC8 function in protein complexes**

While HDAC8 has been traditionally considered to act in the absence of other protein cofactors[6-9, 14, 16-23], a growing body of work suggests that this is incorrect[2, 3, 24-27]. Complexes are important in determining the specificity of other HDACs (e.g. HDAC1, 2, 3), and may be important in regulating HDAC8. While there are a number of putative HDAC8 complexes (Chapter 1) to date, it is unclear whether they are biologically relevant and how they alter biological activity. Furthermore, it is not clear whether there is one large HDAC8-protein complex or many smaller complexes. These complexes may alter HDAC8 localization (i.e. CREB and DEC1), be involved in substrate specificity (like the Sin3b complex is for HDAC1[28]), or be involved in another process such as product release (Chapter 5). The analysis of the substrate binding partners is an important future research topic.



## Bibliography

- [1] H. Lee, N. Sengupta, A. Villagra, N. Rezai-Zadeh, E. Seto, Histone deacetylase 8 safeguards the human ever-shorter telomeres 1B (hEST1B) protein from ubiquitin-mediated degradation, *Mol Cell Biol*, 26 (2006) 5259-5269.
- [2] M. Karolczak-Bayatti, M. Sweeney, J. Cheng, L. Edey, S.C. Robson, S.M. Ulrich, A. Treumann, M.J. Taggart, G.N. Europe-Finner, Acetylation of heat shock protein 20 (Hsp20) regulates human myometrial activity, *J Biol Chem*, 286 (2011) 34346-34355.
- [3] D. Waltregny, W. Glenisson, S.L. Tran, B.J. North, E. Verdin, A. Colige, V. Castronovo, Histone deacetylase HDAC8 associates with smooth muscle alpha-actin and is essential for smooth muscle cell contractility, *FASEB J*, 19 (2005) 966-968.
- [4] M. Haberland, M.H. Mokalled, R.L. Montgomery, E.N. Olson, Epigenetic control of skull morphogenesis by histone deacetylase 8, *Genes Dev*, 23 (2009) 1625-1630.
- [5] I. Van den Wyngaert, W. de Vries, A. Kremer, J. Neefs, P. Verhasselt, W.H. Luyten, S.U. Kass, Cloning and characterization of human histone deacetylase 8, *FEBS Lett*, 478 (2000) 77-83.
- [6] J.J. Buggy, M.L. Sideris, P. Mak, D.D. Lorimer, B. McIntosh, J.M. Clark, Cloning and characterization of a novel human histone deacetylase, HDAC8, *Biochem J*, 350 Pt 1 (2000) 199-205.
- [7] Z.A. Gurard-Levin, M. Mrksich, The activity of HDAC8 depends on local and distal sequences of its peptide substrates, *Biochemistry*, 47 (2008) 6242-6250.
- [8] Z.A. Gurard-Levin, J. Kim, M. Mrksich, Combining mass spectrometry and peptide arrays to profile the specificities of histone deacetylases, *Chembiochem*, 10 (2009) 2159-2161.
- [9] Z.A. Gurard-Levin, K.A. Kilian, J. Kim, K. Bahr, M. Mrksich, Peptide arrays identify isoform-selective substrates for profiling endogenous lysine deacetylase activity, *ACS Chem Biol*, 5 (2010) 863-873.
- [10] B.F. Cravatt, Kinase chemical genomics--a new rule for the exceptions, *Nature methods*, 2 (2005) 411-412.

- [11] M. Krishnamurthy, A. Dugan, A. Nwokoye, Y.H. Fung, J.K. Lancia, C.Y. Majmudar, A.K. Mapp, Caught in the act: covalent cross-linking captures activator-coactivator interactions in vivo, *ACS Chem Biol*, 6 (2011) 1321-1326.
- [12] J.K. Lancia, A. Nwokoye, A. Dugan, C. Joiner, R. Pricer, A.K. Mapp, Sequence context and crosslinking mechanism affect the efficiency of in vivo capture of a protein-protein interaction, *Biopolymers*, 101 (2014) 391-397.
- [13] C.Y. Majmudar, L.W. Lee, J.K. Lancia, A. Nwokoye, Q. Wang, A.M. Wands, L. Wang, A.K. Mapp, Impact of nonnatural amino acid mutagenesis on the in vivo function and binding modes of a transcriptional activator, *J Am Chem Soc*, 131 (2009) 14240-14242.
- [14] D. Riester, C. Hildmann, S. Grunewald, T. Beckers, A. Schwienhorst, Factors affecting the substrate specificity of histone deacetylases, *Biochem Biophys Res Commun*, 357 (2007) 439-445.
- [15] R. Meledin, A. Brik, A. Aharoni, Dissecting the roles of the N- and C-flanking residues of acetyllysine substrates for SIRT1 activity, *Chembiochem*, 14 (2013) 577-581.
- [16] B.E. Schultz, S. Misialek, J. Wu, J. Tang, M.T. Conn, R. Tahilramani, L. Wong, Kinetics and comparative reactivity of human class I and class IIb histone deacetylases, *Biochemistry*, 43 (2004) 11083-11091.
- [17] S. Haider, C.G. Joseph, S. Neidle, C.A. Fierke, M.J. Fuchter, On the function of the internal cavity of histone deacetylase protein 8: R37 is a crucial residue for catalysis, *Bioorg Med Chem Lett*, 21 (2011) 2129-2132.
- [18] D.P. Dowling, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structures of metal-substituted human histone deacetylase 8 provide mechanistic inferences on biological function, *Biochemistry*, 49 (2010) 5048-5056.
- [19] D.P. Dowling, S.L. Gantt, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structural studies of human histone deacetylase 8 and its site-specific variants complexed with substrate and inhibitors, *Biochemistry*, 47 (2008) 13554-13563.
- [20] E. Hu, Z. Chen, T. Fredrickson, Y. Zhu, R. Kirkpatrick, G.F. Zhang, K. Johanson, C.M. Sung, R. Liu, J. Winkler, Cloning and characterization of a novel human class I histone deacetylase that functions as a transcription repressor, *J Biol Chem*, 275 (2000) 15254-15264.
- [21] S.L. Gantt, S.G. Gattis, C.A. Fierke, Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion, *Biochemistry*, 45 (2006) 6170-6178.

[22] S.L. Gantt, C.G. Joseph, C.A. Fierke, Activation and inhibition of histone deacetylase 8 by monovalent cations, *J Biol Chem*, 285 (2010) 6036-6043.

[23] B.C. Smith, J.M. Denu, Acetyl-lysine analog peptides as mechanistic probes of protein deacetylases, *J Biol Chem*, 282 (2007) 37256-37265.

[24] P. Joshi, T.M. Greco, A.J. Guise, Y. Luo, F. Yu, A.I. Nesvizhskii, I.M. Cristea, The functional interactome landscape of the human histone deacetylase family, *Molecular systems biology*, 9 (2013) 672.

[25] G. Canettieri, I. Morantte, E. Guzman, H. Asahara, S. Herzig, S.D. Anderson, J.R. Yates, 3rd, M. Montminy, Attenuation of a phosphorylation-dependent activator by an HDAC-PP1 complex, *Nat Struct Biol*, 10 (2003) 175-181.

[26] Y. Qian, J. Zhang, Y.S. Jung, X. Chen, DEC1 coordinates with HDAC8 to differentially regulate TAp73 and DeltaNp73 expression, *PLoS One*, 9 (2014) e84015.

[27] D. Waltregny, L. De Leval, W. Glenisson, S. Ly Tran, B.J. North, A. Bellahcene, U. Weidle, E. Verdin, V. Castronovo, 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*, 165 (2004) 553-564.

[28] E. Verdin, *Histone deacetylases : transcriptional regulation and other cellular functions*, Humana Press, Totowa, N.J., 2006.

## **Appendices**

## Appendix A

### Kinetics of peptide substrates tested

Corresponding Protein	N-terminus	Sequence	C-terminus	$k_{cat}/K_M$ ( $M^{-1}s^{-1}$ )	$k_{cat}$ ( $s^{-1}$ )	$K_M$
p53	Ac	RHKacKac	Methylcoumarin	800	0.9	1100
H3	Ac	TARKacSTG	NH2	55		
H3	Ac	TKQTARKacSTGGKA	NH2	54		
H3	Ac	ARTKQTARKacSTGGKAP R	NH2	120		
H3	Ac	TGGKacAPR	NH2	5		
H3	Ac	RYQKacSTE	NH2	100		
SWI/SNF complex subunit SMARCC1	Ac	SGKacKGQ	COO-	0.66		
Apoptosis-inducing factor 1	Ac	IIKacDGE	COO-	1.5		
NADH-cytochrome b5 reductase 3	Ac	DIKacYPL	COO-	7.9		
Transcription factor E2F1	Ac	PGKacGVK	COO-	3.7		
Signal peptidase complex subunit 2	Ac	DDKacYTL	COO-	0.67		
Nucleoprotein TPR	Ac	TQKacQEQ	COO-	1.5		
Histone H1.5	Ac	VSKacGTL	COO-	8.5		
E3 ubiquitin-protein ligase TRIM33	Ac	SFKacSDQ	COO-	12		
Histone acetyltransferase KAT6A	Ac	VSKacGPF	COO-	61		
Alpha-enolase	Ac	SGKacYDL	COO-	6.6		

CREB Binding Protein	Ac	DSKacNAK	COO-	5.2		
CREB Binding Protein	Ac	KKKacNNK	COO-	1.7		
Corresponding Protein	N-terminus	Sequence	C-terminus	$k_{cat}/K_M$ ( $M^{-1}s^{-1}$ )	$k_{cat}$ ( $s^{-1}$ )	$K_M$
Histone acetyltransferase KAT6A	Ac	VGKacSVS	COO-	<0.6		
Histone acetyltransferase KAT6A	Ac	DHKacTLY	COO-	7.5		
E3 ubiquitin-protein ligase TRIM33	Ac	FTKacDHL	COO-	3.2		
E3 ubiquitin-protein ligase TRIM33	Ac	KGKacTAQ	COO-	3.1		
Histone H1.5	Ac	GVKacKVA	COO-	1		
Histone H1.5	Ac	KPKacAAK	COO-	4.1		
Histone H1.5	Ac	SFKacLNK	COO-	0.61		
Nucleoprotein TPR	Ac	NQKacLTA	COO-	<0.6		
CAD protein	Ac	SLKacEFY	NH2	78		
DNA mismatch repair protein Msh6	Ac	LSKacFLR	NH2	880		
Cellular tumor antigen p53	Ac	ISKacYDR	NH2	85		
60S ribosomal protein L7	Ac	RKKacGEP	NH2	13		
La-related protein 1	Ac	KIKacRLR	NH2	7.7		
SMC1A	Ac	LGKacFRR	NH2	50		
AMP deaminase 2	Ac	RLKacYSQ	NH2	83		
Interferon regulatory factor 2	Ac	SGKacYFA	NH2	20		
Hematological and neurological expressed 1 protein	Ac	TWKacANF	NH2	77		
Allograft inflammatory factor 1	Ac	TFKacGVD	NH2	14		
Succinate-semialdehyde	Ac	GGKacAFG	NH2	21		

dehydrogenase, mitochondrial						
Actin-binding protein anillin	Ac	FVKacAFA	NH2	2.8		

Corresponding Protein	N-terminus	Sequence	C-terminus	$k_{cat}/K_M$ ( $M^{-1}s^{-1}$ )	$k_{cat}$ ( $s^{-1}$ )	$K_M$
Zinc finger protein 318	Ac	GIKacPFL	NH2	<2		
Elongation factor 1-alpha 1	Ac	FGKacFSW	NH2	4400		
E3 SUMO-protein ligase CBX4	Ac	SFKacYAW	NH2	1200		
DNA-dependent protein kinase catalytic subunit	Ac	SGKacYYY	NH2	21		
Nuclear receptor corepressor 2	Ac	SEKacPEK	NH2	2.6		
SWI/SNF complex subunit SMARCC1	Ac	KGKacDAE	NH2	5.7		
RAI1	Ac	KLGGKacQRAA	NH2	11 ± 1.4	> 0.03	> 3000
MLL2	Ac	SKIQKacQLDQ	NH2	32 ± 5.7	> 0.015	> 400
ZRANB2	Ac	TEIGKacTLAEK	NH2	4.1 ± 0.30	> 0.018	> 4500
SRSF5	Ac	KLSGKacEING	NH2	9.8 ± 1.9	0.018 ± 0.0060	1800 ± 900
THRAP2	Ac	LGDGKacMKS	NH2	4.6 ± 0.10	> 0.015	> 3000
CSRP2BP	Ac	STPVKacFISR	NH2	160 ± 27	0.16 ± 0.040	970 ± 430
ARAD1A	Ac	KLISKacFDKL	NH2	740 ± 36	1.0 ± 0.07	1400 ± 150
NCOA3	Ac	KRILHKacLLQN	NH2	50 ± 5.0	> 0.1	> 1600

## Appendix B

### Protocols:

#### Coupled acetate assay stopped

##### **5x assay buffer**

3 g HEPES (Sigma# H3375)

2 g NaCl (Sigma# 71376)

0.05 g KCl (Sigma# 60128)

pH 8 using metal free NaOH (Sigma# 306576)

Final volume 50 mL add ~2 mL of Chelex resin

##### **Solution 1**

0.48 g HEPES (Fisher# BP310)

0.034 g L-malic acid (Sigma# M1000)

0.21 g NaCl (Sigma# S7653)

0.01 g KCl (Sigma# 60128)

0.4f MgCl (Fisher# BP214)

pH 8 using NaOH

Volume adjust to 32 mL

Steriflip to filter

##### **Solution 2**

36.8 mg ATP (Sigma# A7699)

4 mg CoA (Sigma# C4282)

10 mg NAD<sup>+</sup> (Sigma# 8285)

312  $\mu$ L 5x buffer

1248  $\mu$ L H<sub>2</sub>O

##### **Zn(II)-HDAC8 reconstitution (Co<sup>2+</sup> and Fe<sup>2+</sup> can be substituted in for Zn<sup>2+</sup>):**

###### **Making 100 $\mu$ M Zn:**

6.5  $\mu$ L Zn (1,000 ppm = 15.4  $\mu$ M) (zinc: GFS Chemicals# 1837; cobalt: GFS Chemicals# 1726)

200  $\mu$ L 5x Buffer

793.5  $\mu$ L H<sub>2</sub>O

###### **Reconstitution of Zn(II)-HDAC8: (HDAC8 is reconstituted at 1:1 metal:HDAC8 at 10 $\mu$ M)**



10  $\mu\text{L}$  HDAC8 (657 $\mu\text{M}$  in this case)

131.4  $\mu\text{L}$  5x buffer

449.9  $\mu\text{L}$  H<sub>2</sub>O

65.7  $\mu\text{L}$  Zn (100 $\mu\text{M}$ )

**Acetate Standard Curve:**

**Make 100  $\mu\text{M}$  acetate:**

$\mu\text{L}$  acetate (Ricca Chemical Company# R0091100)

$\mu\text{L}$  5x Buffer

$\mu\text{L}$  H<sub>2</sub>O

**Making standard curve:**

	acetate ( $\mu\text{M}$ )					
	0	1	2	5	10	20
5x buffer	200	200	200	200	200	200
Acetate (100 $\mu\text{M}$ )	0	10	20	50	100	200
H <sub>2</sub> O	800	790	780	750	700	600

Add 60  $\mu\text{L}$  of each standard to a tube containing 2.5  $\mu\text{L}$  10 % HCl and freeze in liquid nitrogen

**Assay:**

Label tubes: 1 for each time point plus one for the reaction tube

Into each tube (except for the reaction tube) put 2.5  $\mu\text{L}$  10 % HCl (diluted in water)

Reaction:

X  $\mu\text{L}$  of peptide

Y  $\mu\text{L}$  of 5x buffer (diluted to 1x final in reaction)

Z  $\mu\text{L}$  of H<sub>2</sub>O (Total-peptide-buffer-HDAC)

A  $\mu\text{L}$  HDAC8 (10  $\mu\text{M}$ )

(Each assay will use 60  $\mu\text{L}$  per time point plus 30  $\mu\text{L}$  on top of that as a buffer so that you don't run out of solution)

Preincubate tubes at 30°C before starting reaction (~10 minutes)

Initiate with metal-HDAC8

To take time points remove ~60 $\mu\text{L}$  of reaction and quench into the 10% HCl

Freeze time points in liquid nitrogen within 20 minutes.

**Coupled reaction: (calculate how much is necessary: ~10  $\mu\text{L}$ /time point + extra)**

600uL Solution 1  
180uL Solution 2  
12uL MDH (Sigma# M2634)  
18uL CS (Sigma# C3260)  
10uL ACS

Mix, spin down and incubate at room temperature for at least 25 minutes before use

1) Unthaw all tubes containing 62.5  $\mu$ L and add 7.5uL of the 6% NaHCO<sub>3</sub> (Mallinckrodt# 7412) to each tube separately. (I've found that this step is not as well done with a multichannel)  
**MAKE SURE TO WEAR EYE PROTECTION DURING THIS STEP.**

2) Vortex and centrifuge each tube

3) transfer 60  $\mu$ L assay and acetate standards from each tube into a 96 well plate that has been cleaned using pressurized air

4) Add 10  $\mu$ L of the coupled reaction mixture to each well using a multichannel pipette

5) Cover the plate (preferably with a lid and foil to protect from light)

6) Incubate plate at room temperature for 45 minutes before reading

7) Read the plate on plate reader at ex = 340nm; em = 460nm

## Histone deacetylation mass spectrometry assay (short protocol)

### 5x assay buffer

3 g HEPES (Sigma# H3375)

2 g NaCl (Sigma# 71376)

0.05 g KCl (Sigma# 60128)

pH 8 using metal free NaOH (Sigma# 306576)

Final volume 50 mL add ~2 mL of Chelex resin

Assay: (STO experiment)

Set up quenches:

Into each quench tube add 1/3 volume of each single time point trichloroacetic acid (e.g. 50  $\mu$ L TCA for a 150  $\mu$ L reaction)

Reaction

X  $\mu$ L HDAC8 (HDAC is varied in this reaction)

Y  $\mu$ L H3/H4 tetramer substrate ( $\sim 1 \times 10^{-10}$  mols/reaction needed)

Z  $\mu$ L 5x Buffer (1x buffer final)

A  $\mu$ L H<sub>2</sub>O (volume needed to make  $\sim 1 \times 10^{-10}$  mols the proper concentration)

For the 0 time point (before adding the HDAC8): Remove 1 time point  
Initiate the reaction with the remaining HDAC8

Take time points by pipetting them into the quench tubes.

Workup:

- 1) Incubate protein solution on ice for 30 minutes
- 2) Centrifuge the samples for 10 minutes at 4°C.
- 3) Remove the supernatant
- 4) Rinse the precipitate with 150  $\mu$ L of -20°C, spin again, and remove the supernatant
- 5) Repeat #4
- 6) Speedvac the pellet until dry.
- 7) Store the pellets at -80°C

Propioylation

- 8) Add 2 $\mu$ L propionic anhydride and 6 $\mu$ L NH<sub>4</sub>OH separately to each tube
- 9) Incubate tubes at 51°C for 1 hour

### Trypsin Digest

- 10) Add 30  $\mu\text{L}$   $\text{NH}_4\text{HCO}_3$  and 2  $\mu\text{L}$   $\text{NH}_4\text{OH}$  to each tube
- 11) pH each tube to (pH 8) using with 2  $\mu\text{L}$  allocations  $\text{NH}_4\text{OH}$
- 12) Once pHed, add 2  $\mu\text{L}$  of .1 mg/mL trypsin to each tube
- 13) Incubate tubes at 37°C overnight

### Mass spec analysis

- 14) Add 3  $\mu\text{L}$  of 10% formic acid to each tube
- 15) Spin tubes for 10 minutes
- 16) Transfer supernatant from tubes into autosampler vials
- 17) Store vials at 4°C until read for use

## Histone deacetylation mass spectrometry assay (long protocol)

### Workup:

- 1) On the day of the assay (after finishing the assay) incubate precipitated protein solution on ice for 30 minutes
- 2) Centrifuge the samples down 10,000-14,000 rpm (benchtop centrifuge) for 10 minutes at 4°C.
  - a. Make sure that the tops of the Eppendorf tubes are pointed the same direction during centrifugation so that you can keep track of the pellet
  - b. After centrifugation the pellets are not always clearly visible
- 3) Remove and discard the supernatant from the tubes
  - a. 100% does not need to be removed, I usually leave 5-20  $\mu\text{L}$  at the bottom of each tube
  - b. Be careful not to disturb the pellet
  - c. If you do disturb the pellet just be careful in subsequent spins not to remove the pellet with the supernatant
- 4) Add 150  $\mu\text{L}$  of -20°C acetone to each tube
- 5) Centrifuge the samples down 10,000-14,000 rpm (benchtop centrifuge) for 10 minutes at 4°C.
- 6) Remove and discard the supernatant from the tubes
- 7) Add 150  $\mu\text{L}$  of -20°C acetone to each tube
- 8) Centrifuge the samples down 10,000-14,000 rpm (benchtop centrifuge) for 10 minutes at 4°C.
- 9) Remove and discard the supernatant from the tubes
- 10) Speedvac the pellet ~10 minutes till dry
- 11) Store pellet at -80°C or proceed immediately

### Propionylation

- 12) Add 2  $\mu\text{L}$  propionic anhydride (Aldrich# 240311) to each tube
  - a. Perform this step in the hood!
  - b. Do not mix in large quantities with  $\text{NH}_4\text{OH}$ !
  - c. Use the propionic anhydride to wet the pellet
- 13) Add 6  $\mu\text{L}$   $\text{NH}_4\text{OH}$  (Aldrich# 33,881-8) to each tube
- 14) Incubate mixture at 51°C for at least 1 hour

### Trypsin Digest

- 15) Add 30  $\mu$ L of 50 mM  $\text{NH}_4\text{HCO}_3$  (Sigma# A6141) and 2  $\mu$ L of  $\text{NH}_4\text{OH}$  to each tube then vortex and spin down each tube (just a few seconds)
- 16) pH each tube to pH ~ 8:
  - a. Dip the tip of a P2 pipette into each tube
    - i. There is no need to suck up any liquid, just use capillary action to get a little
  - b. Place the liquid onto pH paper (pHydralk paper 7-11)
  - c. If the paper turns green or blue (pH  $\leq$  8) the tube is the correct pH, set aside
  - d. If the paper does not turn green or blue add 2  $\mu$ L of  $\text{NH}_4\text{OH}$  then vortex and spin down each tube (just a few seconds). Repeat as necessary
- 17) Once all tubes are the correct add 2  $\mu$ L of 0.1 mg/mL trypsin (Promega# V5113) to each tube
  - a. Trypsin should be resuspended in the buffer that comes with the kit
- 18) Incubate the tubes at 37°C overnight

#### Mass spec analysis

- 19) Add 3  $\mu$ L of 10 % formic acid (Sigma F2004 resuspended in water) into each tube
- 20) Spin tubes for 10 minutes at 10,000-14,000 in the bench top centrifuge
- 21) Transfer the supernatant into a new tube (vial: ThermoScientific National# C014-1. top: ThermoScientific National# C4013-60A. insert: ThermoScientific National# C4012-530)
  - a. Make sure not to get much/any of the precipitate
- 22) Flick vials until the air bubble at the bottom goes away
- 23) Store vials at 4°C until ready to mass spec

## Measuring peptide concentration with Fluorescamine

Create your samples:

- 1) Dilute of peptide 29 (Ac-KGGAKacW-NH<sub>2</sub>) (~5 μM initial concentration) in water:  
peptide/water 1 μL/50 μL, 1 μL /100 μL, 1 μL /150 μL, 1 μL /200 μL
- 2) Dilute desired peptide 1 μL /50 μL, and 1 μL /100 μL
- 3) Measure the OD<sub>280</sub> of the peptide 29 dilutions using the Nanodrop (~2 μL per sample)
- 4) Measure all peptide concentrations and a water sample using the Fluorescamine assay:  
Into a well of a 96 well plate:
  - 75uL Buffer (1M borate pH 9) (Sigma# 7901)
  - 25uL peptide dilution or water sample
  - 15uL of fluorescamine (Acros# 38183-12-9) (15mM in acetone)
  - Mix (~10x up and down) incubate for >10 minutes at room temperature
  - Measure fluorescence ex. 340; em. 460
- 5) Make a standard curve using the  $\epsilon = 5690 \text{ M}^{-1}\text{cm}^{-1}$  for the peptide 29. This can be directly compared to any peptide with one non-acetylated lysine

## HDAC8 purification

### **Expression:**

(expect ~1.5 mg/L protein final)

**Autoinduction TB: (make as much as you need, I typically do a >20 L expression)**

Per 2 liter:

24g Tryptone

48g Yeast Extract

Add 1800mL of H<sub>2</sub>O and autoclave

### **Sugar Mix**

13.8g KH<sub>2</sub>PO<sub>4</sub>

62g K<sub>2</sub>HPO<sub>4</sub>

5mL Glycerol

0.5g Glucose

2g Lactose

Adjust to 600uL H<sub>2</sub>O and sterile filter using a Steritop and Stericup (500mL all will fit)

- 1) Transform HDAC8 B121-DE3 cells and plated onto an LB/AMP plate and incubated at 37C overnight. (Transform HDAC8 fresh each time!)
- 2) In the morning inoculate 5mL 2xYT/AMP using 1 colony and shake at ~200rpm at 37C until the solution is cloudy (~3-5 hours)
- 3) Add the 5mL of 2xYT to 100mL of 2xYT/AMP and shake at 37C until cloudy (~2 hrs)
- 4) Use the culture to inoculate the autoinduction TB  
Into 2L of autoinduction TB the following was added:
  - 100 µg/mL AMP
  - 200 µM ZnSO<sub>4</sub>
  - 25mL of Culture
  - 200mL Sugar mix
- 5) Incubated culture at 30°C overnight shaking ~170rpm
- 6) In the morning pellet the culture using the 6 L centrifuge (5,000 rpm, 10 minutes, 4°C) and stored dry at -80°C or begin the protocol below immediately.

### **Purification:**

**Buffers: (filter all buffers except for MF and PD-10 buffers using a .45uM filters HAWP filter)**



**DEAE low salt buffer (2 L):**

30mM HEPES  
200uM ZnSO<sub>4</sub>  
1mM TCEP  
100mM NaCl  
10ug/mL PMSF  
1ug/mL TAME  
pH 7.8

**DEAE high salt buffer (1 L):**

30 mM HEPES  
200 μM ZnSO<sub>4</sub>  
1 mM TCEP  
500 mM NaCl  
10 μg/mL PMSF  
1 μg/mL TAME  
pH 7.8

**Low imidazole buffer (1 L):**

30 mM HEPES  
100 mM NaCl  
25 mM Imidazole  
1 mM TCEP  
pH 7.8

**High imidazole buffer (1 L):**

30 mM HEPES  
100 mM NaCl  
250 mM Imidazole  
1 mM TCEP  
pH 7.8

**Dialysis buffer 1 (4L):**

25 mM HEPES  
1 mM TCEP  
pH 7.8

**O/N dialysis buffer 1:**

25 mM HEPES  
100 mM NaCl  
1 mM TCEP  
pH 7.8

**Size Exclusion buffer:**

30 mM HEPES  
100 mM NaCl

1 mM TCEP  
pH 7.8

**MF dialysis buffer (w/ EDTA)**

30 mM MOPs (Acros Orgnics# 327665000)  
2 mM KCl (Sigma# 603128)  
1 mM TCEP (Goldbio# TCEP25)  
1 mM EDTA (Sigma# E4884)  
pH 7.8 using MF NaOH

**MF dialysis buffer w/o EDTA**

30 mM MOPs (Acros Orgnics# 327665000)  
2 mM KCl (Sigma# 60128)  
1 mM TCEP (Goldbio# TCEP25)  
pH = 7.8 using MF NaOH

**PD-10 buffer**

25mM HEPES (Sigma# H3375)  
2mM KCl (Sigma# 60128)  
pH = 7.8 using MF NaOH (Sigma# 306576)

- 1) Resuspended cells in ~200mL of DEAE low salt buffer and lysed 2x passes in the microfluidizer
- 2) Put cells into a 200mL beaker with a stir bar on ice and stir
- 3) Add 1% PEI [final] (pH 7.9) relatively slowly (just faster than drop wise) and stirred for 10-15 minutes
- 4) Spin the cell lysate at 18,000rpm for 45 minutes in the big floor centrifuge
- 5) Add the cleared lysate to a 150mL DEAE column equilibrated with 3x volumes of low salt buffer
- 6) Allow the lysate to flow through as fast as it can
- 7) Add 2x column volumes of Low salt buffer and allowed the buffer to flow through as fast as it can by gravity
- 8) Add 200mL of high salt buffer to the column and allowed to flow through as fast can by gravity. (Make sure to collect the elusion!) (There is no need to run a gel in between these steps)
- 9) Put the elusion into dialysis tubing (Spectra/Por MWCO 6-8,00) and incubate in Dialysis buffer 1 for an hour at 4°C

- 10) Mix the dialyzed elution with Imidazole ([final] = 25mM) and 20mL Ni resin
- 11) Allow this to stir on ice for 1 hour
- 12) Add the slurry to an empty column and allowed to flow through as fast as it can by gravity
- 13) Wash the resin with 50mL of 25mM imidazole buffer (50mL low imidazole buffer) and 50mL of 50mM (44 mL Low imidazole buffer 6 mL High imidazole buffer)
- 14) Elute HDAC8 with 50mL of high imidazole buffer
- 15) Run a gel on the purification to confirm the location of the protein
- 16) Add 3 tubes of TEV per 10 L culture to the elution
- 17) Put the TEV/elution in dialysis tubing and dialyze O/N dialysis buffer 1 at 4°C
- 18) In the morning, applied the dialyzed HDAC8 to clean Ni column and collect the FT
- 19) Wash the column with 50mL of 50mM imidazole buffer and collect this wash
- 20) Combine the wash and FT and concentrate in 30mwco filters 3200rpm until the final volume < 3mL.
- 21) Run the protein on the Large Size Hiprep 26/60 Sephacryl S200 column
- 22) Run a gel on the fractions and pool the fractions with pure HDAC8
- 23) Concentrate the fractions to ~0.5mL in 30mwco filters at 3200rpm and put the protein into a <0.5 mL dialysis cassette and dialyze O/N in MF buffer with EDTA at 4°C
- 24) The next day move the cassette to the MF buffer w/o EDTA and incubate O/N at 4°C
- 25) A PD-10 column was run on the HDAC8 using PD-10 buffer and all the fractions containing HDAC8 (check with NanoDrop) were concentrated >500uM and flash frozen in 10uL aliquots

## ACS purification

### **Expression:**

(expect ~1.5 mg/L protein final)

**Autoinduction TB: (make as much as you need, I typically do a >20 L expression)**

Per 2 liter:

24g Tryptone

48g Yeast Extract

Add 1800mL of H<sub>2</sub>O and autoclave

### **Sugar Mix**

13.8g KH<sub>2</sub>PO<sub>4</sub>

62g K<sub>2</sub>HPO<sub>4</sub>

5mL Glycerol

0.5g Glucose

2g Lactose

Adjust to 600uL H<sub>2</sub>O and sterile filter using a Steritop and Stericup (500mL all will fit)

1) Transform ACS plasmid into BI21-DE3 cells and plated onto an LB/AMP plate and incubated at 37C overnight. (Transform HDAC8 fresh each time!)

2) In the morning inoculate 5mL 2xYT/AMP using 1 colony and shake at ~200rpm at 37C until the solution is cloudy (~3-5 hours)

3) Add the 5mL of 2xYT to 100mL of 2xYT/AMP and shake at 37C until cloudy (~2 hrs)

4) Use the culture to inoculate the autoinduction TB

Into 2L of autoinduction TB the following was added:

100 µg/mL AMP

200 µM ZnSO<sub>4</sub>

25mL of Culture

200mL Sugar mix

5) Incubated culture at 30°C overnight shaking ~170rpm

**Buffers: (filter all buffers using a .45uM filters HAWP filter)**

**Low Imidazole Buffer (1 L):**

8.7 g NaCl

7.14 g HEPES

1.36 g Imidazole

0.286 g TCEP

pH 8 using NaOH

Filter solution using a .45uM filters HAWP filter

Store at 4C

**High Imidazole Buffer:**

8.7 g NaCl

7.14 g HEPES

13.6 g Imidazole

0.286 g TCEP

pH 8 using NaOH

Store at 4C

**Size exclusion buffer (1 L):**

8.7 g NaCl

11.9 g HEPES

0.238 g TCEP

pH 8 using NaOH

Store at 4C

**Dialysis Buffer (4L):**

Add a stir bar to the beaker

34.8 g NaCl

28.5 g HEPES

5.44 g Imidazole

1.14 g TCEP

pH 8 using NaOH

Store at 4C

- 6) Spin the cell lysate at 18,000rpm for 45 minutes in the big floor centrifuge
- 7) Load the cleared lysate onto a 10mL Ni-column (either FPLC or gravity)
- 8) Run ~50 mL of low imidazole buffer followed by a ~100 mL gradient of low to high imidazole buffer
- 9) Run a gel to confirm location and purity of ACS
- 10) Pool fractions containing ACS and concentrated in a 30mwco filter to <1mL
- 11) Run the protein on the Large Size Hiprep 26/60 Sephacryl S200 column
- 12) Dialyze with 2 tubes of TEV in against Dialysis buffer
- 13) In the morning run dialyzed ACS on a Ni-tallon column
- 14) Run a gel to confirm the location of ACS and pool ACS

15) (Optional) Assay ACS:

Mix:

100uL solution 1 (see Appendix A: coupled acetate assay)

20uL Solution 2 (see Appendix A: coupled acetate assay)

1uL MDH

.3uL CS

3uL ACS

In a 96 well plate:

Into well 1: 1x buffer

Into well 2: 20uM acetate

Read fluorescence 340nm; em = 460nm

No creep should be observed in well 1 and well 2 should show activity.

16) Concentrate ACS to >750 mM using a 10,000 mwco filter

17) Flash freeze ACS in 10  $\mu$ L fractions and store at -80°C

## Acetylated tetramer purification and assembly

Day 1:

Transform H4 and H3

Make 12L of LB (60g/2L of LB tablets)

Day 2:

O/N growth H3 and H4 (100mL LB each)

Make 10x PBS

80g NaCl

2g KCl

14.4g Na<sub>2</sub>HPO<sub>4</sub>

2.4g KH<sub>2</sub>PO<sub>4</sub>

pH 7.4

Adjust to 1L H<sub>2</sub>O

Make 1x PBS w/20mM nicotinamide

Make 1x PBS w/20mM nicotinamide and 1% (v/v) triton x-100

Day 3:

H3 and H4 expressions

Start cultures 25mL culture /2L LB

For H3:

Grow at 37C

Grow til OD<sub>600</sub>= .5-.7

Add with **1g/L** acetyl-lysine and **2.4g/L** nicotinamide

After 1/2 hour induce with **.119g/L** IPTG

Grow 3 hours and harvest

Resuspend in 1xPBS w/ 20mM Nicotinamide and spin for 10 minutes at 4000rpm

Freeze at -80C

For H4:

Grow at 37C

Grow til OD<sub>600</sub>= .6

Add hour induce with **.119g/L** IPTG

Grow 4 hours and harvest

Freeze at -80C

**Buffers: (No need to filter unless specified)**

**(make buffers containing UREA or Guanidinium chloride day of)**

**Guanidinium chloride Buffer**

114g guanidinium chloride

.5g Tris base

Add a dash of TCEP

pH = 8 - usually no pHing needed

adjust 200mL H<sub>2</sub>O

**His low buffer**

84g Urea

2.4g NaH<sub>2</sub>PO<sub>4</sub>

A dash of TCEP

pH = 6.2

adjust 200mL H<sub>2</sub>O

**His high buffer**

84g Urea

.3g Sodium Acetate

2.3g NaCl

A dash of TCEP

pH = 4.5

adjust 200mL H<sub>2</sub>O

**Q low buffer**

84g Urea

.3g Sodium Acetate

1.1g NaCl

A dash of TCEP

pH = 5.5

adjust 200mL H<sub>2</sub>O

**Q high buffer**

84g Urea

.3g Sodium Acetate

2.2g NaCl

A dash of TCEP

pH = 5.5

adjust 200mL H<sub>2</sub>O

**SP low buffer**



84g Urea  
3g Sodium Acetate  
3.3g NaCl  
A dash of TCEP  
pH = 5.2  
adjust 200mL H2O

**SP high buffer**

84g Urea  
11.5g NaCl  
1.6g Sodium acetate  
.5g Tris  
A dash of TCEP  
pH = 7.5  
adjust 200mL H2O

**Urea dialysis buffer**

840g Urea  
3g Sodium Acetate  
pH = 7.5  
Adjust to 2L of H2O

**H3 TEV buffer:**

24g Tris base  
pH = 7.4  
Adjust to 4L H2O

Day 4 begin purification:

Thaw pellets and resuspend in 1x PBS w/ 20mM nicotinamide with 1 protease inhibitor tablet

Lyse using 2 passes of the microfluidizer

Spin at 18,000 ss34 roter for 15 minutes and decant supernatent

Resuspend cells in PBS + nicotinamide + triton

Spin at 18,000 ss34 roter for 15 minutes and decant supernatent

Resuspend cells in PBS + nicotinamide + triton

Spin at 18,000 ss34 roter for 15 minutes and decant supernatent

Resuspend cells in PBS + nicotinamide

Spin at 18,000 ss34 roter for 15 minutes and decant supernatent

Resuspend cells in PBS + nicotinamide

Spin at 18,000 ss34 roter for 15 minutes and decant supernatent

Macerate pellet in 2mL DMSO for 30 minutes at room temperature

Add 100mL of Guanidinium chloride Buffer and shake at 37C for 1 hr

Spin at 18,000 ss34 roter for 15 minutes and collect the supernatent

For the H4:

Add to 3500mwco dialysis tubing and dialyze in 1L of the Urea dialysis buffer at 4C for the rest of the day

For the H3:

Purify over a 5mL Ni column:

Filter supernatant with a .45uM filter

Put onto AKTA pure and run H3 ni program using the His low buffer and His high buffer

The peak during the gradient is the H3 - collect and put into 3500mwco dialysis tubing into 4L of H2O + 1.4mL BME store at 4C overnight

H4:

Move into a new liter of Urea dialysis buffer at 4C overnight

Day 5:

Place H3 into H3 TEV buffer

Purify H4:

Filter supernatant

Run on the Q column using the Q program in the AKTA pure (remember to set the loop size)  
run a gel on the product

Run an SP column using the SP program in the AKTA pure (remember to set the loop size)  
the peak during the gradient is where the H4 will be found

Put the H4 into 3500mwco dialysis tubing place that in 4L of H2O + 1.5mL BME and incubate at 4C overnight

The H3:

Place 4 tubes of TEV into the H3 and incubate at 4C overnight

Day 6:

Move the H4 and H3 into 4L of H2O + 1.5mL BME in the morning.

Move the H4 and H3 into 4L of H2O + 1.5mL BME around 1.

Make buffers:

**1M HEPES pH = 7.5**

119g HEPES metal free

pH = 7.5 using metal free NaOH

Adjust to 500mL

**If making tetramer:**

**Unfolding buffer**

14.3g Guanidinium HCl

.5mL 1M HEPES pH = 7.5 metal free (Sigma# H3375)

.019g DTT

pH = 7.5

Adjust to 25mL H<sub>2</sub>O

**Refolding buffer with EDTA (make 2x):**

233g NaCl metal free (Sigma# 71376)

20mL 1M HEPES pH = 7.5 metal free (Sigma# H3375)

0.6g EDTA metal free (Sigma# E4884)

1g TCEP (Goldbio# TCEP25)

Adjust to 2L

**Refolding buffer without EDTA and lower NaCl:**

34g NaCl metal free (Sigma# 71376)

20mL 1M HEPES pH = 7.5 metal free

1g TCEP (Goldbio# TCEP25)

Adjust to 2L

Put aside 50mL of the buffer in a metal free falcon tube

At the end of the day ~6pm put the H3 and H4 into tubes and spin

Move the H3 and H4 into new 50mL tubes put ~15mL into each tube and surface freeze in liquid nitrogen

Lyophilize histones

Day 7:

If making tetramer:

Once lyophilized resuspend the H3 and H4 separately in .5mL unfolding buffer and determine the concentration using the nanodrop:

H4:

mw = 15,256

$\epsilon_{276}$  = 5,400

H3:

mw = 11,236

$\epsilon_{276}$  = 4040

Mix the H3 and H4 in a 1:1 molar ratio and dilute to 1mg/mL using the Unfolding buffer.  
Place this mix in 6000-8000 mwco cutoff 23mm flat width tubing  
Place in the Refolding buffer w/ EDTA spinning fast incubate overnight at 4°C

**Day 8:**

Concentrate the sample to <1 mL using a 10,000mwco cutoff filter  
Run the tetramer on the small Size Hiprep 26/60 Sephacryl S200 column using size exclusion buffer

**Size exclusion buffer (1 L):**

116g NaCl  
11.9 g HEPES  
pH 7.5

Run an 4-18% SDS page gel to confirm location of the tetramer

Concentrate the sample to <1 mL using a 10,000mwco cutoff filter

Add the sample into the dialysis tubing or a dialysis cassette 30,000mwco and out in refolding buffer w/ EDTA overnight at 4°C

**Day 9:**

Move the tetramer into refolding buffer w/o EDTA and lower NaCl overnight at 4°C

**Day 10:**

Run a PD-10 column on the tetramer using some fresh refolding buffer w/o EDTA and lower NaCl

Use A<sub>276</sub> to determine the fractions which contain tetramer concentrate the fractions to the desired concentration using a 10,000 or 30,000 mwco spin column

Store protein at 4°C in an ice bucket

## Purifying 601 DNA for nucleosome

**Reference:** Dyer et al. Reconstitution of Nucleosome core Particle, Methods in Enzymology Vol 375, 2004

### General notes:

- plasmid is Amp resistant
- Plasmid has 23 copies of 601 sequence, these are separated by EcoRV sites; there are also 2 EcoRV sites on either side of the insert (25 total EcoRV sites)
- 601 is 147 bp long (total insert is 3.4kbp)
- Best to amplify plasmid in mg quantities and purify using Qiagen Mega or Giga Prep kit (1 Gigaprep is preferred for the amounts of nucleosome we desire)

### General procedure:

1. Digest
2. Precipitate vector
3. check purity of sequence

#### Day 1: Digestion

- 1mL reactions digest plasmid at a concentration of 1mg/ml
- Add 30 units EcoRV per nanomole of EcoRV site in 1x buffer #3 from NEB+1x BSA (10U/reaction)
- Incubate at 37 for 16hr.
- Check completion on 1% agarose gel (1x TBE)
- If not complete, add 50% more enzyme and incubate another 15 hrs.

#### Day2: Precipitate vector

\*\*\*it might be best to test purification with a small amount of your digested product so you see if this protocol works for you or if you have to adjust spin time, % PEG etc.)

- Add 0.192 volume 4M NaCl and 0.346 volume 40% PEG 6000 to digestion (THESE RATIOS ARE CRITICAL SO THAT YOU ONLY PRECIPITATE THE PLASMID BUT NOT THE LOW MW INSERT)
- Incubate on ice for 1hr
- Spin down 27,000g/4 degrees/20min. (or 14000rpm/4degrees/40 min)

- Transfer supernatant to a new tube (601 fragment is here)- run on gel to make sure there is no contamination from vector... if there is contamination titrate in more PEG and repeat spin to get full separation of vector and insert
- Add 2.5 volumes 100% cold ethanol
- Keep at -30 overnight or at -80 for 1hr.
- Spin down 14500rpm/4 degrees/20min
- Carefully remove supernatant; pellet = target DNA
- Air dry DNA by leaving open to air ~10min. (drying completely reduces solubility)
- Dissolve in 2.5mL TE 10/0.1 (or desired volume to get desired concentration) [10mM Tris, 0.1 mM EDTA, pH 8.5]
- probably best to also dialyze into TE to get rid of any residual PEG

If looks good, no need to further purify, store in aliquots of known concentration at -20.