An Unbiased Approach To Identify Endogenous Substrates of “Histone” Deacetylase 8

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

Published Version
doi:10.1021/cb500492r

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:21462463

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
An Unbiased Approach To Identify Endogenous Substrates of “Histone” Deacetylase 8

David E. Olson,† Namrata D. Udeshi,‡ Noah A. Wolfson,† Carol Ann Pitcairn,‡ Eric D. Sullivan,§ Jacob D. Jaffe,† Tanya Svinikina,‡ Ted Natoli,§ Xiaodong Lu,§ Yoshiawa Paulk,‖ Patrick McCarren,‖ Florence F. Wagner,† Doug Barker,† Eleanor Howe,‖ Fanny Lazzaro,‡ Jennifer P. Gale,† Yan-Ling Zhang,† Aravind Subramanian,§ Carol A. Fierke,¶,# Steven A. Carr,‡ and Edward B. Holson*†

†Stanley Center for Psychiatric Research, ¦Proteomics Platform, ©Cancer Program, and ‖Center for the Science of Therapeutics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, United States
‡Department of Biological Chemistry, ¶Interdepartmental Program in Chemical Biology, and #Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States

Supporting Information

ABSTRACT: 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 proteomics methods to identify novel non-histone nuclear substrates of HDAC8, including the tumor suppressor ARID1A. These newly discovered substrates of HDAC8 are involved in diverse biological processes including mitosis, transcription, chromatin remodeling, and RNA splicing and may help guide therapeutic strategies that target the function of HDAC8.

Posttranslational acetylation of lysine residues is a highly conserved and important modification enabling the cellular calibration of protein function and/or stability resulting in effects ranging from cytoskeletal reorganization to changes in gene expression. Histone deacetylases (HDACs) play a key role in maintaining the balance of acetylation states by catalyzing the removal of acetyl groups from the ε-amino groups of acetylated lysine residues. As a result, these enzymes have become important therapeutic targets for a number of disease states including cancer and psychiatric illnesses. As their name implies, HDACs were thought to be primarily responsible for the deacetylation of histones; however, it has become apparent that a large number of non-histone proteins are substrates for these enzymes as well. The HDAC family comprises the NAD(+)-dependent sirtuins (class III) and the metal-dependent HDACs, which can be further divided into three classes (class I: HDACs 1, 2, 3, and 8, class II: HDACs 4, 5, 6, 7, 9, 10, and class IV: HDAC11) based on phylogenetic similarity, with class I being localized primarily in the nucleus and classes II and IV shuttling between the nucleus and the cytoplasm.

Identification of the endogenous substrates of HDAC enzymes is a fundamental area of HDAC research, and this problem has been particularly acute for the class I enzyme HDAC8. Of all the HDACs, HDAC8 is arguably the best characterized structurally. It was the first human class I HDAC structure to be reported, and since then, over 25 additional structures bound to various classes of small molecule ligands and peptides have been disclosed (www.pdb.org). However, despite this knowledge, few of the enzyme’s natural substrates have been identified. To date, only two cellular substrates of HDAC8 have been identified, namely, the estrogen-related receptor alpha (ERR-α) and the structural maintenance of chromosome 3 (SMC3) protein, the latter of which plays a prominent role in Cornelia de Lange syndrome. It remains unclear which, if any, specific histone residues serve as viable substrates for this isoform. In terms of biological function, HDAC8 has been implicated in various cancers including neuroblastoma, uterine, and breast cancer as well as in neural crest development. The HDAC8 substrates that mediate these effects are currently unknown.

To elucidate the cellular substrates and better define the biology of HDAC8, we undertook an unbiased, chemical biology approach that involved monitoring global acetylation and gene expression changes in a representative cell line following treatment with a known, potent, and highly selective small molecule inhibitor of HDAC8. Small molecule modulation coupled with mass spectrometry offers distinct advantages for the identification of acetylation substrates and specific lysine sites responsive to HDAC8 relative to protein knockdown, knockout, or pulldown approaches, including (1) deconvolution of catalytic versus scaffolding functions associated with HDAC8, (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 as well as a suitably designed negative control compound to account for potential compound-driven off-target effects (Figure 1a,b). The inclusion of a negative control compound was particularly important, as PCI-34051 contains a metal-chelating hydroxamic acid group, and this motif has the potential to bind a variety of metalloenzymes. As such, we designed and synthesized BRD3811 (Figure 1a), a compound that retains the hydroxamic acid functionality and contains a minor structural modification to PCI-34051 (i.e., a single methyl group introduced ortho to the hydroxamic acid group) resulting in a 1,000-fold reduction in potency for inhibition of HDAC8 (Figure 1b). Consistent with this finding, molecular docking of PCI-34051 (Figure 1c) and BRD3811 (Figure 1e) into the active site of an HDAC8 crystal structure (PDB accession code 1T64) reveals that the methyl group of BRD3811 cannot be accommodated in the catalytic binding domain of HDAC8 while maintaining an optimal zinc chelation geometry.

Using these chemical tools, we compared the changes in global acetylation in a representative cell line known to express HDAC8 (i.e., MCF7) after treatment with either PCI-34051 (10 μM), BRD3811 (10 μM), or vehicle (DMSO) for 24 h (Figure 2a). Global acetylation profiling was completed by digesting cellular proteins with trypsin and enriching the sequences of our candidate substrates. Our peptide design directed lysine (K) acetylation sites, ensuring that the regulated lysines were flanked on either side by several residues. These “artificial” substrates were incubated with HDAC8, and deacetylation was measured from the production of acetate using an enzyme-coupled assay to determine the steady state kinetic parameters (Figure 2f, Methods). Human Zn-HDAC8 catalyzed the deacetylation of all these synthetic peptides in vitro (Figure 2f), albeit with catalytic efficiencies (kcat/KM) ranging across 3 orders of magnitude. Four of the peptides assayed (i.e., NCOA3, ARID1A, C3SRP2BP, and MLL2) have values of kcat/KM within a factor of 2 or higher than the peptide corresponding to SMC3, providing further evidence that these proteins are likely HDAC8 substrates. In particular, the peptide corresponding to ARID1A is the most efficient non-fluorophore conjugated.

From these data sets, we deemed 5 proteins (i.e., SMC3, RAII, ZRANB2, NCOA3, and THRAP3) to be high-confidence substrates for HDAC8 as they were regulated by 1.5-fold or greater when PCI-34051 treatment was compared 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 ≤0.05) in only one of four experiments.

 Gratifyingly, our unbiased approach successfully identified SMC3, a known substrate of HDAC8, as being significantly regulated by treatment with the HDAC8-selective inhibitor and not BRD3811. Furthermore, we were able to demonstrate that HDAC8-mediated deacetylation occurs on K106 of SMC3, one of two sites known to be acetylated by the acetyltransferase ESCO1. Our coverage of acetylated proteins did not include ERR-α, the only other known cellular HDAC8 substrate, and therefore, we cannot verify its regulation by HDAC8 in MCF7 cells. Our coverage did include several histone proteins (Supplementary Table 1), and we did not observe any significant changes in histone acetylation status upon treatment with PCI-34051 when compared to DMSO or BRD3811. Changes in H2A (H2AFX) and H1.3 (HIST1H1D) acetylation relative to SMC3 are shown for comparison (Figure 2b and c, insets).

To further validate the substrates identified in our proteomics experiments, we devised in vitro enzymatic deacetylation experiments using recombinant human HDAC8 and synthetic acetylated peptides (8–10 aa) based on the sequences of our candidate substrates. Our peptide design centered on the identified lysine (K) acetylation sites, ensuring that the regulated lysines were flanked on either side by several residues. These “artificial” substrates were incubated with HDAC8, and deacetylation was measured from the production of acetate using an enzyme-coupled assay to determine the steady state kinetic parameters (Figure 2f and Methods). Human Zn-HDAC8 catalyzed the deacetylation of all these synthetic peptides in vitro (Figure 2f), albeit with catalytic efficiencies (kcat/KM) ranging across 3 orders of magnitude. Four of the peptides assayed (i.e., NCOA3, ARID1A, C3SRP2BP, and MLL2) have values of kcat/KM within a factor of 2 or higher than the peptide corresponding to SMC3, providing further evidence that these proteins are likely HDAC8 substrates. In particular, the peptide corresponding to ARID1A is the most efficient non-fluorophore conjugated...
peptide substrate of Zn-HDAC8 discovered to date (k_{cat}/K_{M} = 740 M^{-1} s^{-1}) (Figure 2f,g).22 Consistent with previous work, peptides containing an aromatic residue adjacent to the target lysine (e.g., those corresponding to ARID1A and CSRP2BP)
are preferred substrates of HDAC8. To further characterize the enzyme specificity of these synthetic substrates, we profiled the deacetylation activity of commercially available, human recombinant HDACs 1–9 (Supplementary Table 3). While multiple isoforms catalyzed deacetylation of all of the putative substrates, no single peptide was recognized by all of the HDACs tested, and none were uniquely deacetylated by HDAC8. However, the ARID1A peptide exhibited the largest $k_{cat}/K_M$ values for HDAC8 and HDAC8’s closest homologue, HDAC3, with values of 2400 and 2500 M$^{-1}$ s$^{-1}$, respectively (Supplementary Table 3). Finally, we determined the deacetylase inhibitory activity of PCI-34051 toward the ARID1A peptide substrate using commercially available human HDAC8 and calculated a $K_i$ value of 33 nM, demonstrating the ability of this molecule to inhibit HDAC8 in a manner consistent with our cell-based observations (Supplementary Figure 2).

Most of the substrates identified in our study are localized in the nucleus (Figure 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 acetylyme profiling experiments, and this fact was later confirmed in separate, targeted SILAC experiments designed to specifically monitor for histone acetylation changes (Supplementary Figure 3). These results are consistent with previous reports. [23,24]

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 μM (see Methods). We then selected the dose-responsive genes using the IsoGene package [25] (http://CRAN.R-project.org/package=IsoGene). PCI-34051 altered the expression of significantly more genes (70 genes, Supplementary Table 4) than did BRD3811 (11 genes, Supplementary Table 5). While several genes, such as HMOX1, were differentially regulated by both PCI-34051 and BRD3811, the magnitude of the change was much greater for the PCI-34051 treatment.

In an attempt to connect the modulated transcripts to our candidate HDAC8 substrates, we searched for known biological pathway connections between our high-confidence HDAC8 substrates (Figure 2d,e) and the 70 genes dose-responsive to PCI-34051 treatment (Supplementary Table 4) using the Ingenuity IPA Knowledge Base (Figure 3a). No direct associations between HDAC8 and the newly identified substrates were found; however, both the known HDAC8 substrate SMC3 and the tumor suppressor p53 (TP53) were directly connected to HDAC8. Our acetylyme coverage did not include p53, therefore we were unable to determine whether it can be directly modified by HDAC8. Several of the newly identified acetylyme substrates were directly linked to genes differentially expressed upon inhibition of HDAC8. NCOA3, a protein known to positively regulate the expression of HMOX1, is one such example. [26] Additionally, we found 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 increased expression of p21 observed upon treatment with PCI-34051 could be driven in part by the acetylyme changes of ARID1A. Alternatively, increased acetylyme of RAI1 and/or MLL2 could influence the regulation of this important gene. 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. [27] When MCF7 cells are treated with PCI-34051 over the dose range of 0.04–10 μM, a dose-dependent increase in the level of the p21 transcript is observed (Figure 3b). Conversely, BRD3811 increased the level of p21 transcript only at the highest dose tested (i.e., 10 μM) (Figure 3b). To further validate this finding, we examined changes in p21 protein levels after treatment with each compound, and PCI-34051 treatment (10 μM) increased p21 (visualized via Western blot) while BRD3811 (10 μM) treatment did not (Figure 3c,d).

To assess whether our findings extend beyond the context of a single cell type and to 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), HCCS15 (lung), HA1E (kidney), A375 (skin), A549 (lung), and 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 perturbagens using the connectivity map (Cmap) database (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, whereas BRD3811 treatment was not (Supplementary Figure 4). In fact, p21 overexpression was the overexpression perturbation most highly correlated with PCI-34051 treatment, ranking in the 96th percentile on average (BRD3811 treatment did not correlate well with p21 overexpression, 66th percentile rank). It is quite attractive to speculate that some of the anticancer effects of the HDAC8 inhibitor PCI-34051 [19] are mediated in part by increasing p21 levels through these newly discovered substrates. These candidate substrates of HDAC8 may provide a more targeted approach toward specific cancers (or other diseases) driven by the dysregulation of proteins and/or genes within this HDAC8 network. In the case of ovarian clear cell carcinoma, mutations in ARID1A are found in almost half of all cases, and it has been demonstrated that in frame indel mutations fail to induce p21 expression through increased degradation in the nucleus or decreased promoter binding. [28] Efforts toward defining the functional consequences of the change in acetylyme of these proteins by HDAC8 are ongoing.

In conclusion, we have identified several novel substrates of HDAC8 by taking an unbiased approach coupling chemical tools with acetylyme 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 through gene and protein expression studies that inhibition of HDAC8 can affect acetylyme status ultimately influencing 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 the endogenous substrates of other members of the HDAC family of enzymes.

**METHODS**

The details of the methods are described in the Supporting Information.

**ASSOCIATED CONTENT**

1. Supporting Information

   Methods, Supplementary Figures 1–5, Supplementary Tables 1–5, and Supplementary References 26–38. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

*E-mail: edholson@broadinstitute.org.*

**Author Contributions**

We thank M. Weiβer, R. Ratan, S. Sleiman, and A. Desouza for assistance and helpful discussions and L. Gaffney for her help with illustrations. 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 (C.A.F.), T32-GM-008353 (N.A.W.), and T32-GM-008597 (C.A.P.)), and the National Science Foundation (graduate fellowship DGE0718128 (N.A.W.)). This work was supported in part by the Broad Institute of MIT and Harvard and by grants from the US National Cancer Institute (U24CA160034, part of the Clinical Proteomics Tumor Analysis Consortium initiative, to S.A.C.) and National Heart, Lung, and Blood Institute (HHSN268201000033C and R01HL096738 to S.A.C.)

## REFERENCES


**NOTE ADDED AFTER ASAP PUBLICATION**

The surname of contributing author Xiaodong Lu was incorrectly spelled in the version published August 11, 2014. This has been corrected and the revised version was re-posted on August 20, 2014.