

SIRT3 Substrate Specificity Determined by Peptide Arrays and Machine Learning

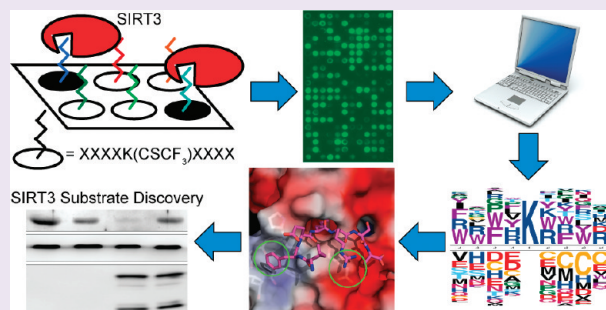
Brian C. Smith,[†] Burr Settles,^{‡,§} William C. Hallows,[†] Mark W. Craven,^{‡,§} and John M. Denu^{†,*}

[†]Department of Biomolecular Chemistry, [‡]Department of Computer Sciences, and [§]Department of Biostatistics & Medical Informatics, University of Wisconsin—Madison, Madison, Wisconsin 53706, United States

S Supporting Information

ABSTRACT: Accumulating evidence suggests that reversible protein acetylation may be a major regulatory mechanism that rivals phosphorylation. With the recent cataloging of thousands of acetylation sites on hundreds of proteins comes the challenge of identifying the acetyltransferases and deacetylases that regulate acetylation levels. Sirtuins are a conserved family of NAD⁺-dependent protein deacetylases that are implicated in genome maintenance, metabolism, cell survival, and lifespan. SIRT3 is the dominant protein deacetylase in mitochondria, and emerging evidence suggests that SIRT3 may control major pathways by deacetylation of central metabolic enzymes. Here, to identify potential SIRT3 substrates, we have developed an

unbiased screening strategy that involves a novel acetyl-lysine analogue (thiotrifluoroacetyl-lysine), SPOT-peptide libraries, machine learning, and kinetic validation. SPOT peptide libraries based on known and potential mitochondrial acetyl-lysine sites were screened for SIRT3 binding and then analyzed using machine learning to establish binding trends. These trends were then applied to the mitochondrial proteome as a whole to predict binding affinity of all lysine sites within human mitochondria. Machine learning prediction of SIRT3 binding correlated with steady-state kinetic k_{cat}/K_m values for 24 acetyl-lysine peptides that possessed a broad range of predicted binding. Thus, SPOT peptide-binding screens and machine learning prediction provides an accurate and efficient method to evaluate sirtuin substrate specificity from a relatively small learning set. These analyses suggest potential SIRT3 substrates involved in several metabolic pathways such as the urea cycle, ATP synthesis, and fatty acid oxidation.



Accumulating evidence suggests that reversible protein acetylation, which was historically reserved for histone proteins, may be a major regulatory mechanism that controls the functions of nonhistone proteins. With the recent cataloging of ~1000 acetylation sites on protein lysine residues^{1–3} comes the challenge of assigning functional roles to specific acetylation sites, identifying the acetyltransferases and deacetylases that regulate acetylation levels, and elucidating the physiological cause and effect of specific lysine acetylation sites. Sirtuins (or Sir2-like proteins) are a conserved family of NAD⁺-dependent protein deacetylases that are implicated in genome maintenance, metabolism, cell survival, and lifespan.^{4,5} The NAD⁺-dependence of the sirtuin reaction suggests that specific protein deacetylation is inextricably linked to metabolism, redox control, and energy status. The observation that the seven mammalian sirtuins (SIRT1–7) display distinct subcellular localization further supports the hypothesis that many, if not most, targets of mammalian sirtuins are nonhistone proteins.

Recent mass spectrometry studies revealed the widespread occurrence of acetylated proteins within mitochondria, as greater than 20% of all mitochondrial proteins are acetylated on at least one lysine residue.^{1–3} Mitochondria are central to cellular metabolism, and mitochondrial dysfunction has been linked to neurodegeneration, diabetes, heart disease, and other age-related diseases.⁶ Understanding the role and function of mitochondrial regulation

by reversible protein acetylation could lead to the better understanding of metabolic function and disease etiology. In humans, SIRT3, SIRT4, and SIRT5 are localized to mitochondria. Among these sirtuins, only SIRT3 displays robust deacetylation activity.⁷ Polymorphisms in human SIRT3 have been linked to survivorship among the elderly,^{8,9} suggesting a possible involvement of SIRT3 in age-related phenomena. The observation that caloric restriction (CR) leads to increases in both transcription and protein levels of SIRT3¹⁰ suggests a positive effect of SIRT3 activity on parameters that influence aging. CR is the only known environmental intervention that strongly increases maximum life span and retards aging in mammals.¹¹ Hints to the role of SIRT3 in metabolic regulation have come from several recent studies. SIRT3 deacetylates and activates mitochondrial acetyl-CoA synthetase 2 (ACS2).^{12,13} Other reports have suggested that SIRT3 deacetylates and activates complex I of the mitochondrial electron transport chain¹⁴ as well as isocitrate, glutamate, and succinate dehydrogenases.^{15,16} SIRT3-deficient mouse embryonic fibroblasts and tissues show lower ATP levels, with basal levels of ATP in the heart, kidney, and liver reduced by >50%.¹⁴ SIRT3 also physically interacts with at least one of the known

Received: July 20, 2010

Accepted: October 14, 2010

Published: October 14, 2010

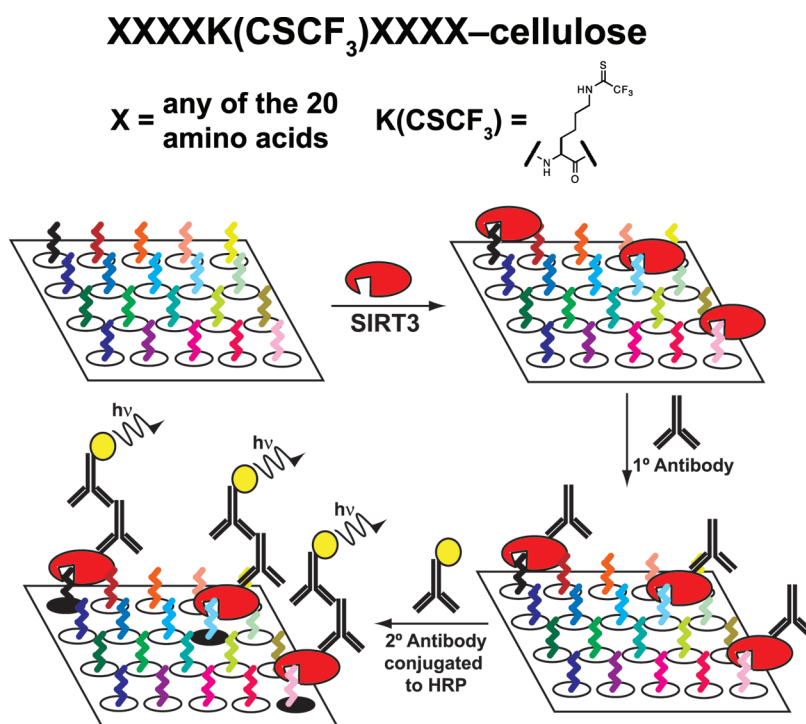


Figure 1. Schematic of SIRT3 binding assay to the SPOT library cellulose membranes. Each individual spot within the SPOT libraries is a 9-mer with four randomized residues proximal to a central thio trifluoroacetyl-lysine residue and C-terminal covalent attachment to amine-modified cellulose membranes.

subunits of Complex I, the 39-kDa protein NDUFA9, and functional studies demonstrate that mitochondria from SIRT3-deficient animals display a selective inhibition of Complex I activity.¹⁴ SIRT3 may modulate mitochondrial function and metabolism in response to metabolic stress. Consistent with this idea, SIRT3 expression in cardiomyocytes is increased under stress, and plays a protective role by preventing Bax-mediated apoptosis.¹⁷ Also, under prolonged fasting, SIRT3-deficient mice display defects in fatty-acid oxidation.¹⁸

To provide an understanding for the molecular basis of SIRT3 function in mitochondria, it is critical that the protein targets and specific lysine deacetylation sites are determined. Unfortunately, in only a few cases have the functional lysines targeted by SIRT3 deacetylation been identified (i.e., acetyl-CoA synthetase 2,^{12,13} cyclophilin D,¹⁹ long-chain acyl coenzyme A dehydrogenase¹⁸). More generally, it will be essential to match functionally important acetylation sites to the enzymes responsible for adding and removing these modifications. Here, to identify potential substrates and understand the pathways controlled by sirtuins, we have developed an unbiased screening strategy that involves a novel acetyl-lysine analogue, SPOT-peptide libraries, machine learning, and kinetic validation. We demonstrate the utility of this novel screening and validation approach using the mitochondrial sirtuin SIRT3.

RESULTS AND DISCUSSION








General Design of SPOT Libraries. To systematically investigate potential sirtuin substrates within a given proteome, we turned to SPOT synthesis,^{20,21} which has recently been utilized. The SPOT technique involves the synthesis of peptides covalently attached via their C-terminus to amine-modified cellulose

membranes (Figure 1). The hydrophilic nature of cellulose minimizes nonspecific binding compared to libraries constructed on hydrophobic polystyrene beads or other surfaces. Because peptides are synthesized in a spatially addressed and parallel manner, SPOT libraries do not require separate sequencing and purification steps to identify queried peptide sequences. The use of SPOT libraries also allows direct quantitation of relative peptide binding affinity.

Here, we utilized SPOT peptide libraries to measure the relative binding affinity of the mitochondrial sirtuin, SIRT3, the major protein deacetylase present in the mitochondria.⁷ Focusing on the mitochondrial proteome of ~1000 proteins²⁵ provided a more manageable data set to explore substrate specificity. The overall strategy is as follows: SIRT3 is incubated with the SPOT membrane and binds preferentially to peptides with the highest affinity (Figure 1). The amount of bound SIRT3, which is proportional to the affinity for each peptide, is then determined utilizing a SIRT3-specific antibody. The SIRT3 primary antibody is detected with HRP-conjugated rabbit IgG secondary antibody. The spots are then visualized, the resulting luminescence is quantified with an imaging camera and software (Figure 1), and the peptide binding preferences are analyzed.

Design of SPOT Libraries: Acetyl-lysine Analogues. The SPOT library design required optimization of several parameters including peptide length, evaluated sequences, and affinity of the central lysine residue. To minimize nonspecific binding and the amount of SIRT3 required for the binding assays, we developed a novel acetyl-lysine analogue that binds to sirtuins tighter than any acetyl-lysine analogue previously tested. Previously, we measured the binding of 10 acetyl-lysine analogue containing peptides to the yeast sirtuin homologue, Hst2, and found that thioacetyl-, trifluoroacetyl-, or propionyl-lysine containing peptides bound

Table 1. Binding constants of various acetyl-lysine analogue peptides to SIRT3 and Hst2

| Acetyl-lysine analog | Structure | SIRT3 K_d (μM) ¹ | Hst2 K_d (μM) ² |
|-----------------------------|---|--|---|
| acetyl-lysine |  | 77 ± 22^3 | 21 ± 4^4 |
| lysine |  | $>10,000^5$ | $>300^5$ |
| thioacetyl-lysine |  | 8.3 ± 0.8^5 | 4.7 ± 1.0^5 |
| trifluoroacetyl-lysine |  | 45.1 ± 8.2^5 | 3.3 ± 0.7^d |
| propionyl-lysine |  | 49.6 ± 13.8^5 | 8.6 ± 0.2^d |
| pentafluoropropionyl-lysine |  | 29.2 ± 5.6^5 | 0.95 ± 0.16^5 |
| thiotrifluoroacetyl-lysine |  | 5.9 ± 0.3^5 | 1.0 ± 0.3^5 |

¹ Determined via competitive fluorescence polarization assays. ² Determined by isothermal titration calorimetry. ³ This work. ⁴ Reference 28. ⁵ Reference 27.

significantly tighter than the corresponding acetyl-lysine containing peptide.²⁶ These findings led us to synthesize the thiotrifluoroacetyl- and pentafluoropropionyl-lysine peptides (Scheme S1; Supplemental Methods) to discover tighter binding acetyl-lysine analogue peptides.

To determine the binding constants of acetyl-lysine analogue containing peptides to SIRT3, a fluorescein-labeled thiotrifluoroacetyl-lysine peptide based on the peptide sequence of the known SIRT3 substrate acetyl-CoA synthetase 2 (fluorescein-LPKTRSGK(CSCF₃)VMRR-OH; Fluor-ACS2) was synthesized.^{12,13} Fluorescence polarization measurements yielded a K_d value of $1.7 \pm 0.3 \mu\text{M}$ for direct binding of Fluor-ACS2 to SIRT3, which was similar to the K_d value of $0.9 \pm 0.2 \mu\text{M}$ determined by isothermal titration calorimetry (ITC). Subsequently, using Fluor-ACS2 in a competitive fluorescence polarization assay, the K_d values were determined for various acetyl-lysine analogue peptides based on the human histone H3 sequence modified at Lys-14 (H₂N-KSTGGK(acetyl analogue)APRKQ-OH) (Supplemental Figure S1, Table 1). To establish binding trends, the yeast sirtuin Hst2 was also evaluated^{26–28} (Table 1, Supplemental Figure S1).

The thioamide and trifluoromethyl group within the acetyl-lysine analogues provided independent but additive binding effects. The thioacetyl- and trifluoroacetyl-lysine peptide bound SIRT3 9.3-fold and 1.7-fold tighter than the corresponding acetyl-lysine peptide (Table 1, Supplemental Figure S2). Therefore, we expected the thiotrifluoroacetyl-lysine peptide would bind 15.8-fold ($K_d \approx 4.9 \mu\text{M}$) tighter to SIRT3 compared to the acetyl-lysine peptide. Indeed, we observed a K_d value of $5.9 \mu\text{M}$ for thiotrifluoroacetyl-lysine peptide binding to SIRT3. Using the same analysis for Hst2, the thiotrifluoroacetyl-lysine

peptide was predicted to bind 28.4-fold tighter ($K_d \approx 0.7 \mu\text{M}$) to Hst2, and a K_d value of $1.0 \mu\text{M}$ was observed, suggesting thiotrifluoroacetyl-lysine peptides are general tight-binding sirtuin probes.

Fluorination of the propionyl group to form the pentafluoropropionyl analogue also behaved as predicted for both SIRT3 and Hst2 binding. In particular, the propionyl- and trifluoroacetyl-lysine peptides bound SIRT3 1.55-fold and 1.7-fold tighter than the corresponding acetyl-lysine peptide. Therefore, we predicted that the pentafluoropropionyl-lysine peptide would bind 2.65-fold tighter ($K_d \approx 29.1 \mu\text{M}$) to SIRT3 compared to the acetyl-lysine peptide, and we observed a K_d value of $29.2 \mu\text{M}$ (Table 1, Supplemental Figure S2). For Hst2, pentafluoropropionyl-lysine peptide was predicted to bind 15.5-fold ($K_d \approx 1.35 \mu\text{M}$) tighter to Hst2, and we observed a K_d value of $0.95 \mu\text{M}$. Unlike the thiotrifluoroacetyl analogue, which exhibited similar fold improvements in SIRT3 and Hst2, the pentafluoropropionyl analogue exhibited a significantly greater improvement in Hst2 binding compared to SIRT3 binding. We recently reported that propionyl- and acetyl-lysine peptides exhibited similar overall turnover rates with Hst2; however, significantly slower turnover rates were observed for propionyl- compared to acetyl-lysine peptides with SIRT3.²⁶ Therefore, the pentafluoropropionyl moiety is likely well accommodated in the Hst2 active site but not in the SIRT3 active site, whereas the thiotrifluoroacetyl moiety is well accommodated in both active sites. As the thiotrifluoroacetyl-lysine analogue displayed general tight binding to both SIRT3 and Hst2, this analogue was used in subsequent SPOT library screens.

Design of SPOT Library: Peptide Length. Another important consideration in the design of the SPOT libraries was the length of peptides. For maximal effectiveness of the method, the peptides should be short enough to ensure synthetic integrity but long enough to achieve high binding affinity and discrimination among library members. To this end, we synthesized control SPOT libraries containing thiotrifluoroacetyl-lysine surrounded by peptide sequences based on the general *in vitro* sirtuin substrate lysine-14 of histone H3,^{26,29} the SIRT1 substrate p53,^{30–33} the SIRT2 substrate α -tubulin,³² the SIRT3 substrate acetyl-CoA synthetase 2,^{12,13} and the SIRT6 substrate lysine-9 of histone H3.³⁴ In these control libraries, peptides longer than nine amino acids did not result in increased SIRT3 binding (Supplemental Figure S3).

We also examined published sirtuin crystal structures in which acetyl-lysine containing peptides were bound in the acetyl-lysine binding pocket.^{35–43} In these structures, ordered residues were only reported for up to four residues N- or C-terminal to acetyl-lysine despite using peptides of length 8–18 amino acids in the crystallization conditions. This suggests that the majority of specific interactions in the sirtuin acetyl-lysine substrate binding pocket occur within the four residues immediately N- and C-terminal to acetyl-lysine.

A recent bioinformatics study reported acetylation motifs from the computational analysis of existing proteomic acetyl-lysine data sets.⁴⁴ Although seven N- and C-terminal residues were utilized in training data sets, only the immediate N-terminal residue and four C-terminal residues revealed motifs that were statistically correlated with acetylation.

Finally, the only previous example of machine learning analysis of SPOT libraries, which identified novel antimicrobial peptides, utilized 9-mer peptides.^{45,46} For the above four reasons, we utilized 9-mer peptides in our SPOT libraries with four amino

acids N- and C-terminal to the central thiotrifluoroacetyl-lysine residue (Figure 1).

Design of SPOT Library: Peptide Sequences. As there are 25.6 billion possible sequences for the 20 natural amino acids at eight randomized positions, screening every possible sequence was impractical. As SIRT3 is localized to the mitochondria,^{47–50} the sequence space can be narrowed significantly by focusing on the mitochondrial proteome. Even this yields $\sim 22,000$ possible 9-mer sequences with a central lysine. Therefore, to provide the greatest coverage of the relevant sequence space, we utilized two distinct libraries. One SPOT library focused on known acetylation sites within the mitochondria and contained the human sequences of acetylated peptides from a recent acetyl proteomics screen (229 peptides).¹ As all acetyl proteomic studies have relied on the ability of anti acetyl-lysine antibodies to enrich for acetylated peptides in tryptic digests, known acetylation sites are potentially biased toward both the specificity of antibodies as well as the length of peptides required for efficient mass spectral sequencing. To minimize any bias, the other SPOT library contained 300 randomly sampled lysine-centered 9-mers from the mitochondrial proteome, independent of evidence for acetylation.

Screening of SPOT Libraries for SIRT3 binding. After identification and validation of the proper acetyl-lysine analogue, peptide length, and peptide sequences to be utilized, we screened the resulting SPOT libraries for their ability to bind SIRT3. Each library was synthesized in triplicate, screened with SIRT3 and the resulting integrated optical density (IOD) binding scores were averaged to reduce measurement noise and improve the training signal for the machine learning model (see below). A representative SIRT3 library screen of 300 randomly sampled lysine-centered 9-mers from the mitochondrial proteome is shown in Figure 2. These screens exhibited a wide range of binding affinity from high intensity spots representing high affinity peptides to spots with background signal intensity. The identical procedure performed in the absence of SIRT3 revealed no binding above background for both libraries (data not shown), which indicated that the observed SIRT3 binding to the membrane in Figure 2 was not the result of nondirected antibody binding.

Machine Learning of SIRT3 Binding Affinity. As mentioned above, testing all 22,000 mitochondrial lysine centered 9-mer sequences was impractical, so we developed a method to model and predict SIRT3 binding for sequences not examined in the SPOT libraries. Machine learning was recently used in the identification of novel antibacterial peptides from SPOT libraries.^{45,46} To model binding affinity for peptide sequences, we used a linear regression approach that was trained using the IOD scores from our SPOT library screens and predicted IOD scores for other peptides not present in our libraries. This method (described below) induces a model by treating a peptide's IOD score as a linear function of various molecular properties. It performed as well or better than other machine learning algorithms we investigated (e.g., artificial neural networks, Gaussian processes, and decision trees). Linear regression has the added advantage of easily interpretable model parameters (i.e., the coefficients of the learned function).

To represent peptides for the machine learning algorithm, the amino acids at the eight variable positions of each peptide were represented as binary variables. For example, the sequence IINQKRFND has value 1 for Ile at the -4 position and Arg at the $+1$ position, and value 0 for all other amino acids at those positions, resulting in $8 \times 20 = 160$ total individual residue features. Additionally, we used 10 global chemical properties to describe

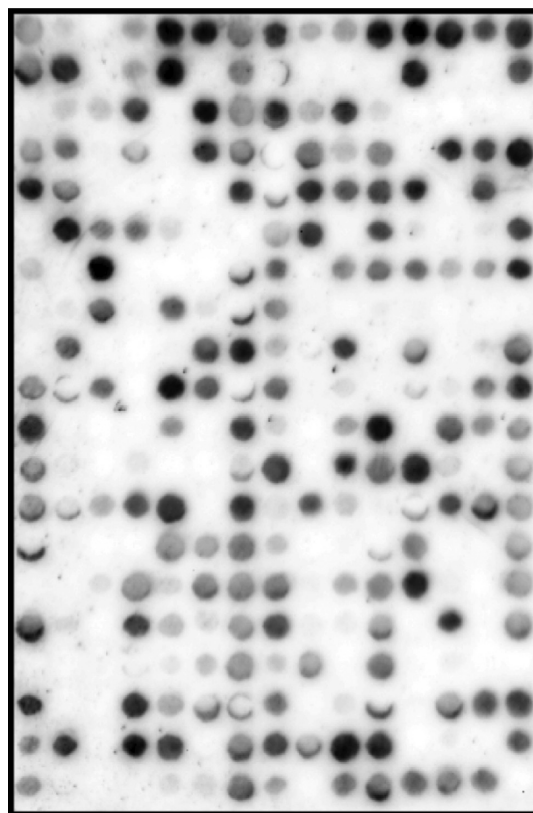


Figure 2. Binding of SIRT3 to a SPOT library of 300 lysine-centered 9-mer peptides randomly selected from the mitochondrial proteome. A representative SPOT library is shown with SIRT3. Performing the same procedure without SIRT3 gave no signal above background (data not shown). SPOT library screens were performed as described under Methods.

each peptide. These global properties were hydrophobicity;⁵¹ flexibility;⁵² conformational tendencies for α helix, β strand, β turn, and coil;⁵³ average area buried on transfer to folded protein;⁵⁴ and accessibility.⁵⁵ We desired to separate the conformational preferences for parallel and antiparallel β strands, so we also included a separate scale for these preferences.⁵⁶ These scales provide values for individual amino acids, and thus we took the average value across all residues in a sequence to produce a real-numbered value for each peptide as a whole. We then normalized the input value for each scale feature to be in the range from 0 to 1. This simple normalization allowed us to better compare the relative contribution of each scale's parameter weight in the analysis of our learned model.

To fit the parameters of our model, we used least-squares regression. However, we found that the common approach of computing the coefficients in closed form led to significant “overfitting,” which occurs when a statistical model represents spurious correlations and noise in the training signal rather than an actual underlying relationship between the input and output variables. This sometimes occurs in machine learning when the model has too many degrees of freedom relative to the amount of training data available, which is likely the case in our study with 170 input features and 529 training sequences.

To remedy this, we implemented a gradient-based numerical optimization approach based on the L-BFGS algorithm,⁵⁷ which finds the least-squares solution iteratively, and we employed “early stopping” to prevent overfitting.⁵⁸ Specifically, we halt the training process once the reduction in estimated squared error

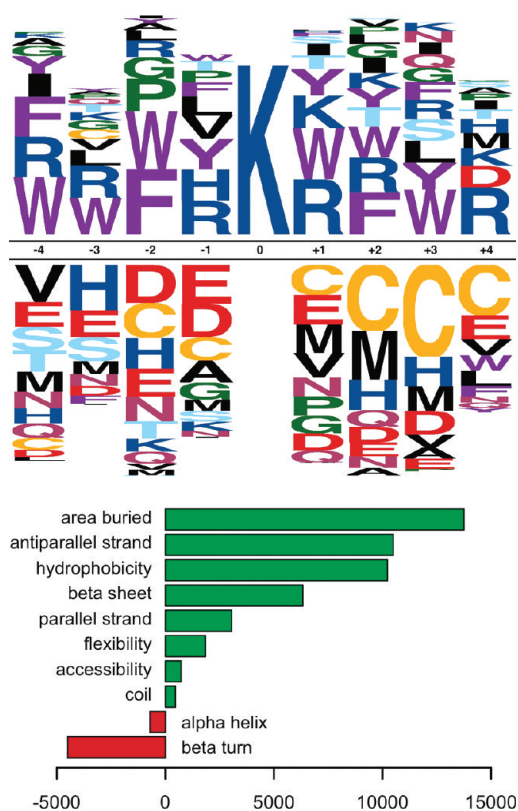


Figure 3. (Top) Sequence logos illustrating learned linear regression weights for the position-specific residue features. The height of each residue is proportional to the coefficient used by the model in predicting SIRT3 binding specificity. Positively correlated residue positions are shown above the baseline, and negatively correlated residues are below. (Bottom) Global feature model parameters. A bar chart illustrating learned linear regression weights for the global peptide scale features. Positive values suggest that the corresponding chemical property (e.g., hydrophobicity) or secondary structure (e.g., antiparallel strand) is positively correlated with SIRT3 binding specificity. Negative values likewise suggest a negatively correlated residues.

falls below 0.001. One could also use other forms of parameter regularization⁵⁹ to maintain predictive accuracy on novel peptides, but we found early stopping to be sufficient. A more detailed discussion of this method, and a comparison to other machine learning approaches we investigated will be discussed elsewhere (manuscript in preparation). We also plan to publicly release a software tool that implements our feature generation and model training for the research community at large.

Parameter weights for the trained machine learning model of SIRT3 substrate specificity are presented in Figure 3. The sequence logo representation in Figure 3, top panel illustrates the relative contribution of residues at specific positions surrounding the central lysine to SIRT3 substrate specificity. Residues above the baseline suggest positive correlation with SIRT3 binding specificity, whereas residues below the baseline suggest negatively correlated relationships, and larger characters suggest greater relative impact on binding. We emphasize that these coefficients do not reflect relative frequency of each residue in the libraries (which is the interpretation for some other sequence logos in the literature), but rather how discriminative they are in predicting high or low binding affinity, in terms of IOD score. The bar chart in Figure 3, bottom panel illustrates the analogous relationship between global parameters and SIRT3 substrate

specificity. Again, positive weights suggest positive correlation with measured SIRT3 binding affinity, and negative weights indicate negatively correlated relationships, and the magnitude of each weight indicates its relative impact on SIRT3 binding.

To evaluate the predictive accuracy of our model, we conducted a cross-validation study using the 529 peptide measurements from our SPOT libraries. For each peptide, this involved training a model using the remaining 528 sequences and predicting a binding score for the single held-out peptide. This procedure is repeated for every sequence in the library, allowing us to fairly estimate how well the model can predict binding on the SPOT membrane for sequences for which it was not trained. The Pearson correlation between predicted scores and the actual IOD measurements is $R^2 = 0.54$, indicating that our model predictions can explain over half of the observed variance in actual IOD measurements from SPOT array screenings.

Machine Learning Prediction of SIRT3 Substrate Specificity Across the Mitochondrial Proteome. A significant advantage of our machine learning model is the ability to predict SIRT3 substrate specificity for sequences not present among the original 529 peptides tested in our SPOT libraries. To this end, we used the MitoCarta human inventory of ~ 1000 mitochondrial proteins²⁵ to generate a list of $\sim 22,000$ lysine-centered 9-mers as potential SIRT3 binding sites. We then generated residue and global scale features for these peptide sequences and predicted their binding affinities using the learned model described above. We have included these predictions in Supplemental Table 2. This table has also been annotated with known acetylation sites from the acetyl-lysine proteomics by Kim et al., Choudhary et al., and Zhao et al.^{1–3} Since the acetylation sites discovered in these proteomics screens likely represent a subset of the acetyl-lysine sites present in mitochondria, we have also annotated the spreadsheet with the result of an online acetylation prediction algorithm (phosida.com).

Validation of SPOT Library Screening and Machine Learning. Implicit in the machine learning model is the hypothesis that binding correlates with substrate specificity in deacetylation assays. This may not be the case as some peptides may bind outside the active site and potentially uncover other relevant protein:protein interaction sites. However, we have purposely tried to minimize the contribution of binding outside the active site through engineering a tight binding acetyl-lysine analogue (thiotrifluoroacetyl-lysine) that should result in increased binding to only the SIRT3 active site and not other potential binding sites.

To evaluate the results from the SPOT libraries and the machine learning prediction of SIRT3 substrate specificity, we synthesized the acetyl-lysine versions of sequences with a wide variety of both experimentally tested (11 peptides) and predicted (13 peptides) binding affinities and measured their steady-state kinetic k_{cat}/K_m values with SIRT3 using our previously published continuous sirtuin assay.⁶⁰ While not the most sensitive sirtuin assay, this assay was an excellent choice to rapidly evaluate the steady-state k_{cat}/K_m values among a large diverse group of peptide substrates for SIRT3. The measured k_{cat}/K_m values varied ~ 300 -fold over the range of peptides tested. To evaluate the correlation of the results, we plotted our *measured* IOD scores versus of the $\log k_{cat}/K_m$ values.^{61,62} The IOD scores reflect the relative binding affinity of SIRT3 toward each thiotrifluoroacetylated SPOT peptide. The k_{cat}/K_m parameter is the apparent second order rate constant between SIRT3 and free, acetylated peptide. This steady-state constant reflects both the affinity for substrate and

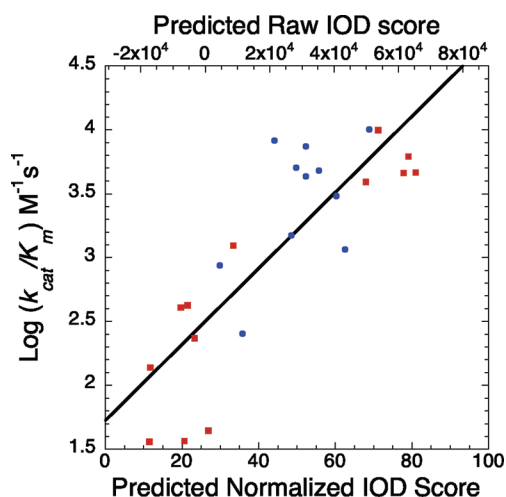


Figure 4. Steady state kinetic validation of acetyl-lysine peptides selected from SPOT-library screens of SIRT3. Predicted percentile refers to the relative binding affinity predicted from the learned model trained on the binding intensities from SPOT library screens (Figure 2). Blue circles are sequences tested on the SPOT membrane, and red squares represent sequences that were not tested.

the efficiency of deacetylation. Therefore, evaluation of the correlation between IOD scores and k_{cat}/K_m values was critical as the SPOT libraries may have selected for peptides that bind to the active site in a conformation that did not allow for efficient deacetylation. This analysis revealed a fairly strong linear correlation (R^2 value of 0.37, $p = 0.047$) of binding affinity of thio-trifluoroacetyl-lysine containing peptides on SPOT membranes and deacetylation efficiency of acetyl-lysine containing peptides in solution for the 11 peptides we analyzed. To further evaluate our machine learning model, we also plotted the *predicted* IOD scores from the machine learning model versus the $\log k_{\text{cat}}/K_m$ values (Figure 4). This analysis revealed a slightly weaker correlation for the 11 peptides in the experimentally tested set (R^2 value of 0.30, $p = 0.079$) but an excellent correlation for all 24 of the peptides we analyzed *in vitro* (R^2 value of 0.59, $p < 0.001$), which suggested that the machine learning model is accurate in predicting deacetylation efficiency for peptides not experimentally tested on the SPOT membranes. Collectively, these analyses provided strong evidence that peptide-binding affinity on membranes revealed in the SPOT library binding screens correlated with deacetylation efficiency in solution.

Substrate Specificity and Mitochondrial Sirtuins. The combined approach of SPOT library screening and machine learning analysis revealed several interesting determinants for efficient SIRT3-catalyzed deacetylation. In general, aromatic residues (i.e., Phe, Trp, and Tyr) exhibited tight SIRT3 binding. Another general trend was negatively charged residues (i.e., Asp and Glu) were disfavored across the entire 9-mer peptide (Figure 3, top panel). Cysteine residues were generally disfavored, especially C-terminal to the acetyl-lysine residue. This may be due to negative charge resulting from cysteine deprotonation or oxidation to sulfinic/sulfonic acids. Alternatively, these cysteines may form disulfides between peptides and limit SIRT3 accessibility within the SPOT membrane. Likewise, positively charged residues (i.e., Arg and Lys) were correlated with tight SIRT3 binding. This is consistent with the recently published SIRT3 structure⁶³ that reveals an overall negative charge throughout the peptide binding pocket, particularly at

residues C-terminal to acetyl-lysine (Figure 5, Supplemental Figure S5a).

A previous acetyl-lysine proteomics study revealed that acetyl-lysine sites are enriched in regions with ordered secondary structure, in particular α -helical and β -sheet regions.² Here, we found that peptides with a high propensity to form α helix or β -turn structures were disfavored as SIRT3 substrates; however, peptides with a high propensity to form β -strand structures were favored overall (Figure 3, bottom panel). In particular, peptides with a high propensity to form antiparallel-strands were highly favored whereas parallel-strands were slightly favored. This observation is highly consistent with all known sirtuin crystal structures in which acetyl-lysine peptides bind in an extended antiparallel-strand conformation.^{35–43} Within the SIRT3 active site, E296, P297, and L398 of SIRT3 form an antiparallel β strand with the -2 and -1 position and acetyl-lysine of the substrate peptide, and E323, V324, and E325 of SIRT3 form an antiparallel β strand with the $+1$, $+2$, and $+3$ positions of the substrate peptide (Supplemental Figure S4). In addition, peptide flexibility was less favored toward SIRT3 binding when compared to β -strand structures. This finding is in contrast to a previous study that reported the yeast sirtuin, Hst2, prefers conformationally flexible substrates within unstructured protein regions.⁶⁴

In general, acetyl-lysine residues reside in regions with higher average solvent accessibility than the average of all lysines,¹ suggesting that acetyl-lysine residues occur in more exposed environments. However, our machine learning analysis indicated that SIRT3 preferentially binds regions with lower accessibility and higher area-buried scores, reflecting residues that would prefer to be buried given their inherent chemical nature. It is important to note that peptides with low accessibility and high area-buried scores are not necessarily less solvent accessible in specific proteins. As binding to SIRT3 provides a surface to bury these residues, this may provide additional binding energy, analogous to the hydrophobic effect in protein folding.⁶⁵

A recent bioinformatics study by Schwartz et al. analyzed known acetyl-lysine sites in mammals and found several motifs that were highly correlated with lysine acetylation.⁴⁴ Interestingly, several of these motifs were correlated with SIRT3 binding, indicating that peptides predicted to be SIRT3 substrates are predicted to be acetylated. The acetylation motifs K(ac)K, K(ac)R, K(ac)Y, K(ac)XF, and K(ac)XXXX were highly correlated with SIRT3 binding, whereas the acetylation motifs K(ac)XXK and K(ac)F were less highly correlated with SIRT3 binding (Figure 3, top panel). The only acetylation motif negatively correlated with SIRT3 binding was GK(ac); this motif likely represents nuclear histone acetylation (the GK motif)^{2,66} and would not be expected to correlate with SIRT3 substrate specificity in the mitochondria.

Recall that the 229-peptide library has a potential sample bias as it contains only known acetylated peptides in the mitochondria. To ensure that the observed relationship between the Schwartz et al. motifs and SIRT3 binding was not an artifact of over-representation in our training set, we compared the frequency of each motif in the acetylation library, the randomized library, and the observed frequencies among the full set of $\sim 22,000$ lysine-centered 9-mers in the mitochondrion (the background set). Only the K(ac)Y motif appeared at a significantly higher frequency in the acetylated library than in the randomized library or the background set ($p < 0.05$, using binomial tests with Bonferroni correction, details in Supplemental Table 1). However, the over-representation of the K(ac)Y

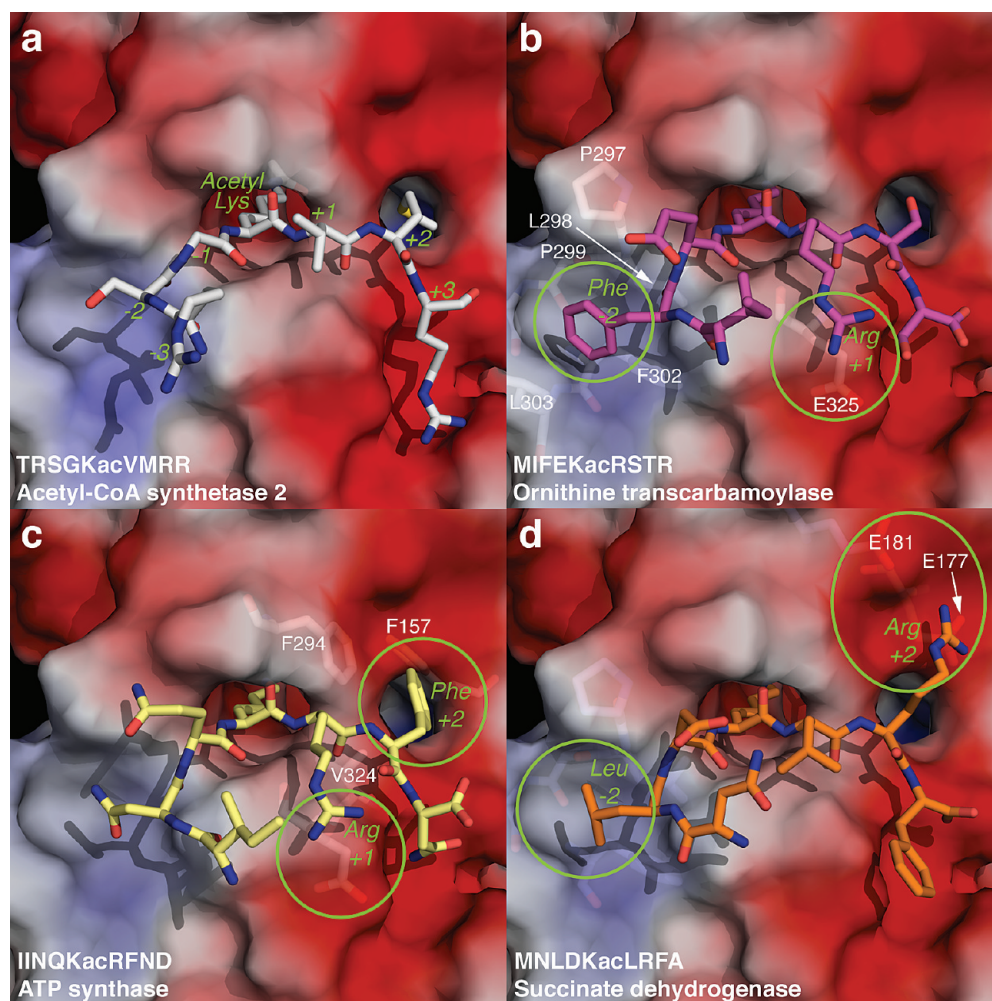


Figure 5. Models of acetyl-lysine peptides bound in the SIRT3 active site. Stick models of (a) acetyl-CoA synthetase 2 (ACS2), (b) K88 of ornithine transcarbamoylase (OTC), (c) K230 of the α subunit of ATP synthase, and (d) K498 of the flavoprotein subunit of succinate dehydrogenase docked into the SIRT3 active site were based on the structure of a thioalkylamidate bound to SIRT3 resulting from the reaction of a thioacetyl-lysine acetyl-CoA synthetase 2 (ACS2) peptide and NAD^+ (PDB code 3GLT).⁶³ The enzyme surface shown is prior to energy minimization and colored on the basis of electrostatic potential with red, blue, and white surface representing negative, positive, and hydrophobic residues, respectively. Modeling was performed by mutating an acetyl-lysine ACS2 peptide bound to SIRT3 (PDB code 3GLR⁶³) using Pymol,⁷⁵ and the structure was energy-minimized using CHARMMing⁷⁶ using the CHARMM defaults for Shake. Images were generated using Pymol.⁷⁵

motif likely does not impact our findings, since (i) the randomized library has virtually the exact same distribution of K(ac)Y as the full set ($p = 1.0$), which helps alleviate any bias, and (ii) other motifs such as K(ac)K and K(ac)R show even stronger correlation with SIRT3 binding while not being over-represented.

To separate acetylation motifs by cellular compartment (mitochondria, nucleus, and cytoplasm), a more recent acetyl-lysine proteomics study analyzed separately the relative abundance of amino acids flanking acetyl-lysine residues.² Interestingly, Choudhary et al. found that Phe was enriched at the -2 position next to acetyl-lysine residues in the mitochondria, which matches the preference of SIRT3 at this position (Figure 3, top panel). Additionally, Phe and Tyr were enriched at the $+2$ position for acetyl-lysine residues; SIRT3 has a dual preference for positively charged (i.e., Lys and Arg) in addition to aromatic residues (i.e., Phe, Tyr, Trp) at this position. Similarly, Tyr is enriched at the $+1$ position next to mitochondrial acetyl-lysine residues, and SIRT3 prefers Lys and Arg at the $+1$ position but possesses a

secondary preference for aromatic residues such as Trp and Tyr. However, Choudhary et al. found that the residues flanking many of the identified mitochondrial acetyl-lysine sites are enriched for negatively charged residues; this is divergent from the preferences of SIRT3 and suggests that this negatively charged subset of mitochondrial acetyl-lysine sites may instead be SIRT4 and/or SIRT5 substrates. Comparing the active sites of SIRT3 and SIRT4 supports this hypothesis. The peptide binding site in SIRT3 is negatively charged, whereas the same site in SIRT4 is positively charged (Supplemental Figure S5). However, the peptide binding site in SIRT5 is mostly hydrophobic, suggesting that SIRT3, SIRT4, and SIRT5 possess very distinct protein substrate preferences.

Although we made a concerted effort to construct and utilize the SPOT libraries to provide an efficient method for rapidly screening sirtuin substrate specificity, the ability of this approach to provide insight into cellular targets is limited to sirtuins that discriminate potential substrates on the basis of the amino acid sequence surrounding the acetyl-lysine. To date, the

contribution of interacting surfaces outside the active site between sirtuins and their substrate targets is completely unknown. There are no available costructures with native acetylated proteins. Consequently, it is unclear if there are quaternary-structure contributions to specificity that would not necessarily be revealed in the SPOT library analysis presented here. In addition, it is unknown if the induced backbone conformation of the peptide when present in a whole protein substrate plays a role in substrate specificity. Furthermore, the overall inaccessibility (i.e., buried within tertiary structure) of a particular lysine may restrict its ability to be acetylated by acetyltransferases and subsequently deacetylated by sirtuins. Therefore, additional information about lysine accessibility within a protein substrate could help eliminate unlikely targets that would otherwise be predicted to be excellent SIRT3 substrates (e.g., lysine sites with high area buried/hydrophobicity and low accessibility scores).

Modeling of SIRT3 with Potential Substrates. We next sought to rationalize our learned model of SIRT3 peptide substrate specificity in the context of a recently reported SIRT3 crystal structure bound with an acetyl-CoA synthetase 2 (ACS2) peptide.⁶³ The combined total heights of the residues depicted in Figure 3, top panel provide insight into the relative importance of each position to SIRT3 binding. In particular, the -4 , -2 , $+1$, $+2$, and $+3$ positions provide the greatest contributions toward SIRT3 binding. Therefore, we modeled several peptide sequences into the SIRT3 peptide binding site, and analyzed the -2 , $+1$, and $+2$ positions of the peptide for interactions with SIRT3. The -4 position was not ordered and the side chain conformation at the $+3$ position varied significantly in the SIRT3 structures and therefore we were unable to analyze specific interactions at these positions. We were unable to reliably model the side chain conformations at the -1 position due to the presence of a glycine at this position in the ACS2 peptide that was used for modeling. However, as mentioned above, our learned model indicates that the -1 position is not as important in dictating substrate specificity as the -4 , -2 , $+1$, $+2$, and $+3$ positions. To explore structural features at these positions, the ACS2 peptide (Figure 5a) was mutated to three separate peptides, all of which exhibited tight binding in our SPOT libraries: K88 of ornithine transcarbamoylase (OTC), K230 of the α subunit of ATP synthase, and K498 of the flavoprotein subunit of succinate dehydrogenase. These modeled peptides were then energy minimized within the SIRT3 structure and examined for salient interactions. It is important to note that these models are necessarily qualitative and strongly suggest the need for further structural work with alternate peptide sequences bound to SIRT3.

For the OTC peptide, the interactions at two residues were particularly striking, Phe at the -2 position and Arg at the $+1$ position (Figure 5b). Phe was the most highly favored residue at the -2 position in the machine learning analysis, and Phe docked into a hydrophobic cavity on the SIRT3 surface defined by P297, L298, P299, F302, and L303. Arg was the most highly favored residue at the $+1$ position and this Arg modeled close to the negatively charged E325 on the surface of SIRT3.

The interactions at the $+1$ and $+2$ position were of interest for the ATP synthase peptide (Figure 5c). Similar to the OTC peptide, ATP synthase also contains an Arg at the $+1$ position that may interact with E325 of SIRT3. SIRT3 possesses a hydrophobic hole near the $+2$ position of the peptide defined by F157, F294, and V324. Phe is the most highly favored residue at the $+2$ position in the learned model (Figure 4a),

and this hole may accommodate hydrophobic residues such as Phe.

The succinate dehydrogenase peptide also contains two interactions of interest (Figure 5d). Similar to the OTC peptide, the succinate dehydrogenase peptide contains a hydrophobic Leu residue at the -2 position (Phe in both OTC and SCHAD) that may dock in the same hydrophobic cavity. The succinate dehydrogenase peptide contains an Arg residue at the $+2$ position. Arg is the second most highly favored residue at the $+2$ position (Figure 3a). This Arg may extend beyond the hydrophobic hole at the $+2$ position to instead interact with the negatively charged E177 or E181. This may explain the dual preference for positively charged and hydrophobic residues at the $+2$ position in the learned model.

Identifying SIRT3 Substrates. One main extension of this study is to accelerate the ability to identify sirtuin substrates. In this section, we will highlight the utility of this method using several examples. A potential SIRT3 substrate identified from our SPOT library screens is ornithine transcarbamoylase (OTC). OTC is a key enzyme in the urea cycle that converts ornithine and carbamoyl phosphate into citrulline. Lysine 88 of OTC ranked in the top 8% of the SPOT library screen of known mitochondrial acetylation sites (Supplemental Table 1). In addition, our machine learning predictions ranked K88 of OTC first among the 24 lysines present in the OTC sequence and the top 4% of all mitochondrial peptides (Supplemental Table 2). Interestingly, a recent study revealed that OTC K88 acetylation decreases the affinity for carbamoyl phosphate and the maximum velocity.⁶⁷ Our SPOT libraries and machine learning analysis strongly suggested that SIRT3 is the enzyme that deacetylates K88 of OTC *in vivo*. Indeed, in a companion study, we demonstrated *in vivo* and *in vitro* that OTC is a *bona fide* target of SIRT3 and that SIRT3 modulates the urea cycle through direct OTC deacetylation (manuscript to be submitted).

Other relevant proteins whose corresponding peptides exhibited tight binding to SIRT3 are subunits of ATP synthase and the electron transport chain (NADH dehydrogenase). Indeed, the tightest binding peptide in the SPOT library based on known mitochondrial acetylation sites was from ATP synthase, H1 transporting, mitochondrial F1 complex, α subunit 1 (H1NQK(ac)-RFND) (Supplemental Table 1). Interestingly, both the α and β subunits of ATP synthase were identified as SIRT3 interacting partners in a recent proteomics screen.⁶⁸ The authors also verified the interaction between the α subunit of ATP synthase and SIRT3 by coimmunoprecipitation, but did not examine deacetylation. Previous studies have shown that ATP levels and mitochondrial electron potential are decreased in the *sirt3*^{-/-} mouse.¹⁴ A potential explanation for these lowered ATP levels is that hyperacetylation of ATP synthase upon loss of SIRT3 inhibits ATP synthase and electron transport chain activity.

Several peptides from enzymes involved in fatty acid oxidation exhibited tight binding in our SPOT library screens with SIRT3. Acetylated peptides from short chain L-3-hydroxyacyl-Coenzyme A dehydrogenase (SCHAD) (K202 in top 6% of known mitochondrial acetylation SPOT library and top 3% of all mitochondrial peptides in the learned model), very-long-chain acyl-CoA synthetase (VLACS) (K291 in top 13% of known mitochondrial acetylation SPOT library), and 3-ketoacyl-Coenzyme A thiolase (K137 in top 12% of mitochondrial acetylation SPOT library and top 1% of all mitochondrial peptides in the learned model) scored highly as potential substrates for SIRT3 deacetylation

(Supplemental Tables 1 and 2). Consistent with an important role in β -oxidation, *sirt3*^{-/-} mice exhibit abnormally high levels of a number of acylcarnitines (ref 18 and manuscript to be submitted). Recently, it was reported that long-chain acyl coenzyme A dehydrogenase (LCAD) is regulated by acetylation and deacetylation by SIRT3.¹⁸

Identifying SIRT3 substrates is paramount to understanding how reversible acetylation can modulate mitochondrial metabolism, apoptosis, redox, and potentially lifespan extension. This study provides a compelling insight into possible SIRT3 substrates in the mitochondria. We anticipate that this data set will be used to discover novel SIRT3 substrates in one of two ways: (i) A discovery mode where biologists search among previously cataloged acetylation sites that are functionally uncharacterized and predicted to be SIRT3 substrates by our machine learning analysis. This list of potential SIRT3 substrates can be further narrowed by other factors such as location of the lysine within the protein (e.g., within the active site of an enzyme or mitochondrial location sequence that is cleaved upon translocation) or location of the protein within the mitochondria (SIRT3, SIRT4, and SIRT5 are present within the mitochondrial matrix). (ii) A validation mode where a biologist has discovered a potential SIRT3 substrate, but the site of deacetylation is unknown. In this case, our learned model would prioritize the lysines as potential sites of SIRT3-catalyzed deacetylation. In this case, a customized SPOT library could be constructed containing only sequences surrounding the lysines within the protein of interest and tested for SIRT3 binding. Overall, this new approach will greatly expand the scope and quicken the pace of sirtuin substrate identification.

METHODS

General Materials and Methods. SIRT3 and anti-goat-HRP conjugated antibody were purchased from Abcam (Cambridge, MA, USA). Fmoc-protected amino acids were purchased from peptide international (Louisville, KY, USA). All other chemicals used were of the highest purity commercially available and were purchased from Sigma (St. Louis, MO, USA), Aldrich (Milwaukee, WI, USA), or Fisher Scientific (Pittsburgh, PA, USA). Mass spectral analyses were performed at the University of Wisconsin-Madison Biotechnology Center mass spectrometry facility.

Sirtuin Expression and Purification. Hst2^{69,70} and SIRT3¹² was expressed recombinantly in *E. coli* and purified using Ni-NTA affinity resin as described. The enzyme concentrations were determined using the method of Bradford using BSA as the standard.⁷¹

SIRT3 Peptide Binding Constants by Fluorescence Anisotropy. SIRT3 (10 μ M) was incubated with 20 nM Fluor-ACS2 peptide (see Supplemental Experimental Procedures for synthesis) and various concentrations of unlabeled acetyl-lysine analogue peptides ranging from 100 nM to 10 mM in 50 mM Tris-HCl (pH 7.5) and 10% (v/v) glycerol in 100 μ L total volume. Fluorescence anisotropy (mA) was measured at 25 $^{\circ}$ C using a Panvera Beacon 2000 FP system with 490 nm excitation and 530 nm emission wavelengths. Anisotropy values at each competitor peptide concentration were converted to fraction Fluor-ACS2 bound (F_{SB}) using 1eq 1:

$$F_{SB} = \frac{A_{OBS} - A_F}{(A_B - A_{OBS})Q + A_{OBS} - A_F} \quad (1)$$

where A_{OBS} is the measured anisotropy at the particular competitor peptide concentration, A_F is the anisotropy of free Fluor-ACS2, A_B is the anisotropy of the fully SIRT3-bound Fluor-ACS2, and Q is the ratio of fluorescence intensities of bound and free Fluor-ACS2. IC_{50} values were

then determined from plots of fraction bound versus competitor peptide concentration using 2eq 2:

$$F_{SB} = F_{SB}^0 \left(1 - \frac{[\text{peptide}]}{IC_{50} + [\text{peptide}]} \right) \quad (2)$$

where F_{SB}^0 is the fraction Fluor-ACS2 bound with no added competitor peptide, and $[\text{peptide}]$ is the concentration of added competitor peptide. Competitor peptide equilibrium dissociation constants were then calculated as previously described.^{72,73}

Synthesis of SPOT Library Membranes. The SPOT peptide libraries were synthesized on an Intavis Respep XL automated peptide synthesis robot. Libraries were synthesized on commercially available amino-functionalized cellulose sheets (INTAVIS; Köln, Germany) according to the manufacturers protocols using standard Fmoc peptide synthesis techniques (for general references of SPOT synthesis, see refs 20 and 21). Double coupling was used for the first seven amino acids, and triple coupling for each subsequent amino acid. Thiotrifluoroacetyl-lysine was added to peptides via use of the Fmoc-protected derivative (Fmoc-Lys(CSCF₃)-OH; see Supplemental Methods). Using the Respep XL robot, up to 600 peptides can be synthesized per 10 cm \times 15 cm sheet. After synthesis the side-chain protecting groups were cleaved using 95% v/v TFA, 0.3% v/v triisopropylsilane, and 4.7% v/v H₂O. Membranes were then washed with dichloromethane (4 \times 20 mL), DMF (2 \times 10 mL), water (2 \times 10 mL), and ethanol (2 \times 10 mL) and air-dried. Successful synthesis was monitored visually by tryptophan fluorescence under a UV lamp at 360 nm.

SIRT3 Binding to SPOT Libraries. SPOT library membrane was incubated overnight at 4 $^{\circ}$ C with 5% w/v BSA in PBST 0.05%. The membrane was incubated with 2 μ M of SIRT3 in PBS for 1 h and then with 5% w/v BSA in PBST 0.05% for 30 min. The membrane was then incubated with a primary rabbit antibody specific for SIRT3 for 1 h in 5% w/v BSA in PBST 0.05%, washed (3 \times 5 min with PBST 0.05%), incubated with a goat anti-rabbit IgG-HRP conjugated secondary antibody, washed (3 \times 5 min with PBST 0.05%), and developed with Supersignal West Dura from Thermo Scientific. Images of the developed SPOT membranes were captured utilizing Lab Works 4.6 software in a Epi Chem II Darkroom by UVP, and spot intensity was integrated to give integrated optical density (IOD) scores.

Machine Learning. A linear regression model was trained using a least-squares fit to the IOD measurements obtained via SPOT library screens. In particular, we used a combination of input features that (i) describe the residues flanking the central lysine and (ii) capture global properties of the peptide, such as hydrophobicity and secondary structure information. The induced model provides some insight into the chemical and structural properties that characterize peptides with high SIRT3 binding specificity. The model was then used to predict binding affinity scores for over 22,000 potential SIRT3 binding sites in the mitochondrial proteome. Full details will be described in a subsequent publication (manuscript in preparation).

Solid-Phase Peptide Synthesis of In Vitro Validation Peptides. Acetyl-lysine containing peptides for *in vitro* validation of our SPOT libraries were synthesized on a 5 μ mol scale on an Intavis Respep XL automated peptide synthesis robot according to the manufacturers protocols using standard tBu/Fmoc solid-phase peptide synthesis techniques. Acetyl-lysine was added to peptides via Fmoc-Lys-(COCH₃)-OH. The protecting groups used were Boc for lysine and tryptophan, *tert*-butyl for serine and threonine, Pbf for arginine, Trt for cysteine, and tBu for glutamine, asparagine, glutamate, aspartate, and tyrosine. After completion of the synthesis, the resin was rinsed with dichloromethane and dried. The full-length peptide was then deprotected and cleaved from the resin with 95% v/v TFA, 0.3% v/v triisopropylsilane, and 4.7% v/v H₂O. The peptide was precipitated with cold ethyl ether and sedimented by centrifuge washing twice with

additional cold ethyl ether. The precipitate was dried, redissolved in water, and lyophilized. Masses of the cleaved peptides were confirmed by MALDI-TOF mass spectrometry. Crude peptides were purified by semipreparative RP-HPLC on a C18 small pore column (Grace Vydac, 10 mm × 250 mm, 10 μm) eluting with a gradient of 0–100% acetonitrile with 0.02% TFA in water with 0.05% TFA. Fractions collected were lyophilized to yield final peptides as dry white powders.

Determination of Peptide Concentrations. Acetyl-lysine peptide concentrations were determined from the change in NADH absorbance at 340 nm using sirtuin enzyme-coupled assays.⁶⁰ Typical assay mixtures contained ~100 μM acetyl-lysine peptide, 0.2 mM NAD⁺, 0.2 mM NADH, 1 mM DTT, 3.3 mM α-ketoglutarate, 1 μM MBP-PncA (nicotinamidase), 3 units of glutamate dehydrogenase from bovine liver, and ~1 μM Hst2 in 20 mM potassium phosphate at pH 7.5. Reactions were run to completion, and the peptide concentration was calculated from the change in NADH absorbance using an extinction coefficient of 6.22 mM⁻¹ cm⁻¹ and a path length of 0.9 cm for 300 μL reactions. Concentrations of peptides containing tyrosine and/or tryptophan were independently determined using calculated extinction coefficients at 280 nm as described previously.⁷⁴ These two methods routinely gave peptide concentrations within 20% of each other.

Determination of SIRT3 Kinetic Parameters. The k_{cat} and K_m values for different acetyl-lysine containing peptides were determined using a previously described sirtuin enzyme-coupled assay.⁶⁰ Typical assay mixtures contained 10–600 μM acetyl-lysine peptide, 600 μM NAD⁺, 0.2 mM NADH, 1 mM DTT, 3.3 mM α-ketoglutarate, 1 μM MBP-PncA (nicotinamidase), 3 units of glutamate dehydrogenase from bovine liver, and 0.5–1 μM SIRT3 in 20 mM potassium phosphate at pH 7.5. Reactions were carried out in a final volume of 300 μL per well in a clear, flat-bottomed, 96-well plate. All assay components except SIRT3 were preincubated at 25 °C until absorbance at 340 nm stabilized, and the reaction was initiated by the addition of SIRT3. The rates were analyzed continuously by measuring NADH consumption at 340 nm and were determined from the slopes of the initial linear portion of each curve up to 20% conversion using a path length of 0.9 cm. The background rates of reactions lacking SIRT3 resulting from the spontaneous formation of nicotinamide or ammonia were subtracted from the initial velocities of the SIRT3-catalyzed reactions. The platereader used for our assay possessed a resolution of 0.001, which corresponds to a detection limit of ~0.2 μM.

■ ASSOCIATED CONTENT

Supporting Information. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*jmdenu@wisc.edu.

■ ACKNOWLEDGMENT

We thank Mark Devries for assistance in SPOT library synthesis and *in vitro* analysis of SIRT3 substrates and Sam Oliver for help acquiring mass spectra of peptides. This work was supported by a grant from the National Institutes of Health (GM065386). B.S. was supported by a grant from the National Human Genome Research Institute (T32HG002760).

■ REFERENCES

(1) Kim, S. C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L., Grishin, N. V., White, M., Yang, X. J., and

Zhao, Y. (2006) Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell* 23, 607–618.

(2) Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834–840.

(3) Zhao, S., Xu, W., Jiang, W., Yu, W., Lin, Y., Zhang, T., Yao, J., Zhou, L., Zeng, Y., Li, H., Li, Y., Shi, J., An, W., Hancock, S. M., He, F., Qin, L., Chin, J., Yang, P., Chen, X., Lei, Q., Xiong, Y., and Guan, K. L. (2010) Regulation of cellular metabolism by protein lysine acetylation. *Science* 327, 1000–1004.

(4) Imai, S., and Guarente, L. (2010) Ten years of NAD-dependent SIR2 family deacetylases: implications for metabolic diseases. *Trends Pharmacol. Sci.* 31, 212–220.

(5) Longo, V. D., and Kennedy, B. K. (2006) Sirtuins in aging and age-related disease. *Cell* 126, 257–268.

(6) Lenaz, G., Baracca, A., Fato, R., Genova, M. L., and Solaini, G. (2006) New insights into structure and function of mitochondria and their role in aging and disease. *Antioxid. Redox Signaling* 8, 417–437.

(7) Lombard, D. B., Alt, F. W., Cheng, H. L., Bunkenborg, J., Streeper, R. S., Mostoslavsky, R., Kim, J., Yancopoulos, G., Valenzuela, D., Murphy, A., Yang, Y., Chen, Y., Hirsche, M. D., Bronson, R. T., Haigis, M., Guarente, L. P., Farese, R. V., Weissman, S., Verdin, E., and Schwer, B. (2007) Mammalian Sir2 homolog SIRT3 regulates global mitochondrial lysine acetylation. *Mol. Cell. Biol.* 27, 8807–8814.

(8) Rose, G., Dato, S., Altomare, K., Bellizzi, D., Garasto, S., Greco, V., Passarino, G., Feraco, E., Mari, V., Barbi, C., BonaFe, M., Franceschi, C., Tan, Q., Boiko, S., Yashin, A. I., and De Benedictis, G. (2003) Variability of the SIRT3 gene, human silent information regulator Sir2 homologue, and survivorship in the elderly. *Exp. Gerontol.* 38, 1065–1070.

(9) Bellizzi, D., Rose, G., Cavalcante, P., Covello, G., Dato, S., De Rango, F., Greco, V., Maggiolini, M., Feraco, E., Mari, V., Franceschi, C., Passarino, G., and De Benedictis, G. (2005) A novel VNTR enhancer within the SIRT3 gene, a human homologue of SIR2, is associated with survival at oldest ages. *Genomics* 85, 258–263.

(10) Shi, T., Wang, F., Stieren, E., and Tong, Q. (2005) SIRT3, a mitochondrial sirtuin deacetylase, regulates mitochondrial function and thermogenesis in brown adipocytes. *J. Biol. Chem.* 280, 13560–13567.

(11) Weindruch, R., Walford, R. W. (1988) *The Retardation of Aging and Disease by Dietary Restriction*, C. C. Thomas, Springfield, IL.

(12) Hallows, W. C., Lee, S., and Denu, J. M. (2006) Sirtuins deacetylate and activate mammalian acetyl-CoA synthetases. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10230–10235.

(13) Schwer, B., Bunkenborg, J., Verdin, R. O., Andersen, J. S., and Verdin, E. (2006) Reversible lysine acetylation controls the activity of the mitochondrial enzyme acetyl-CoA synthetase 2. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10224–10229.

(14) Ahn, B. H., Kim, H. S., Song, S., Lee, I. H., Liu, J., Vassilopoulos, A., Deng, C. X., and Finkel, T. (2008) A role for the mitochondrial deacetylase Sirt3 in regulating energy homeostasis. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14447–14452.

(15) Schlicker, C., Gertz, M., Papatheodorou, P., Kachholz, B., Becker, C. F., and Steegborn, C. (2008) Substrates and regulation mechanisms for the human mitochondrial sirtuins sirt3 and sirt5. *J. Mol. Biol.* 382, 790–801.

(16) Cimen, H., Han, M. J., Yang, Y., Tong, Q., Koc, H., and Koc, E. C. (2010) Regulation of succinate dehydrogenase activity by SIRT3 in mammalian mitochondria. *Biochemistry* 49, 304–311.

(17) Sundaresan, N. R., Samant, S. A., Pillai, V. B., Rajamohan, S. B., and Gupta, M. P. (2008) SIRT3 is a stress-responsive deacetylase in cardiomyocytes that protects cells from stress-mediated cell death by deacetylation of Ku70. *Mol. Cell. Biol.* 28, 6384–6401.

(18) Hirsche, M. D., Shimazu, T., Goetzman, E., Jing, E., Schwer, B., Lombard, D. B., Grueter, C. A., Harris, C., Biddinger, S., Ilkayeva, O. R., Stevens, R. D., Li, Y., Saha, A. K., Ruderman, N. B., Bain, J. R., Newgard, C. B., Farese, R. V., Alt, F. W., Kahn, C. R., and Verdin, E. (2010) SIRT3 regulates mitochondrial fatty-acid oxidation by reversible enzyme deacetylation. *Nature* 464, 121–125.

- (19) Shulga, N., Wilson-Smith, R., and Pastorino, J. G. (2010) Sirtuin-3 deacetylation of cyclophilin D induces dissociation of hexokinase II from the mitochondria. *J. Cell Sci.* 123, 894–902.
- (20) Frank, R. (2002) The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports—principles and applications. *J. Immunol. Methods* 267, 13–26.
- (21) Hilpert, K., Winkler, D. F., and Hancock, R. E. (2007) Peptide arrays on cellulose support: SPOT synthesis, a time and cost efficient method for synthesis of large numbers of peptides in a parallel and addressable fashion. *Nat. Protoc.* 2, 1333–1349.
- (22) Rathert, P., Dhayalan, A., Ma, H., and Jeltsch, A. (2008) Specificity of protein lysine methyltransferases and methods for detection of lysine methylation of non-histone proteins. *Mol. Biosyst.* 4, 1186–1190.
- (23) Rathert, P., Dhayalan, A., Murakami, M., Zhang, X., Tamas, R., Jurkowska, R., Komatsu, Y., Shinkai, Y., Cheng, X., and Jeltsch, A. (2008) Protein lysine methyltransferase G9a acts on non-histone targets. *Nat. Chem. Biol.* 4, 344–346.
- (24) Rathert, P., Zhang, X., Freund, C., Cheng, X., and Jeltsch, A. (2008) Analysis of the substrate specificity of the Dim-5 histone lysine methyltransferase using peptide arrays. *Chem. Biol.* 15, 5–11.
- (25) Pagliarini, D. J., Calvo, S. E., Chang, B., Sheth, S. A., Vafai, S. B., Ong, S. E., Walford, G. A., Sugiana, C., Boneh, A., Chen, W. K., Hill, D. E., Vidal, M., Evans, J. G., Thorburn, D. R., Carr, S. A., and Mootha, V. K. (2008) A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 134, 112–123.
- (26) Smith, B. C., and Denu, J. M. (2007) Acetyl-lysine analog peptides as mechanistic probes of protein deacetylases. *J. Biol. Chem.* 282, 37256–37265.
- (27) Smith, B. C., and Denu, J. M. (2007) Mechanism-based inhibition of sir2 deacetylases by thioacetyl-lysine Peptide. *Biochemistry* 46, 14478–14486.
- (28) Smith, B. C., and Denu, J. M. (2007) Sir2 deacetylases exhibit nucleophilic participation of acetyl-lysine in NAD⁺ cleavage. *J. Am. Chem. Soc.* 129, 5802–5803.
- (29) Borra, M. T., Langer, M. R., Slama, J. T., and Denu, J. M. (2004) Substrate specificity and kinetic mechanism of the Sir2 family of NAD⁺-dependent histone/protein deacetylases. *Biochemistry* 43, 9877–9887.
- (30) Langley, E., Pearson, M., Faretta, M., Bauer, U. M., Frye, R. A., Minucci, S., Pelicci, P. G., and Kouzarides, T. (2002) Human SIR2 deacetylates p53 and antagonizes PML/p53-induced cellular senescence. *Embo J.* 21, 2383–2396.
- (31) Luo, J., Nikolaev, A. Y., Imai, S., Chen, D., Su, F., Shiloh, A., Guarente, L., and Gu, W. (2001) Negative control of p53 by Sir2alpha promotes cell survival under stress. *Cell* 107, 137–148.
- (32) North, B. J., Marshall, B. L., Borra, M. T., Denu, J. M., and Verdin, E. (2003) The human Sir2 ortholog, SIRT2, is an NAD⁺-dependent tubulin deacetylase. *Mol. Cell* 11, 437–444.
- (33) Vaziri, H., Dessain, S. K., Ng Eaton, E., Imai, S. I., Frye, R. A., Pandita, T. K., Guarente, L., and Weinberg, R. A. (2001) hSIR2(SIRT1) functions as an NAD-dependent p53 deacetylase. *Cell* 107, 149–159.
- (34) Michishita, E., McCord, R. A., Berber, E., Kioi, M., Padilla-Nash, H., Damian, M., Cheung, P., Kusumoto, R., Kawahara, T. L., Barrett, J. C., Chang, H. Y., Bohr, V. A., Ried, T., Gozani, O., and Chua, K. F. (2008) SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. *Nature* 452, 492–496.
- (35) Avalos, J. L., Bever, K. M., and Wolberger, C. (2005) Mechanism of sirtuin inhibition by nicotinamide: altering the NAD⁺ cosubstrate specificity of a Sir2 enzyme. *Mol. Cell* 17, 855–868.
- (36) Avalos, J. L., Celic, I., Muhammad, S., Cosgrove, M. S., Boeke, J. D., and Wolberger, C. (2002) Structure of a Sir2 enzyme bound to an acetylated p53 peptide. *Mol. Cell* 10, 523–535.
- (37) Cosgrove, M. S., Bever, K., Avalos, J. L., Muhammad, S., Zhang, X., and Wolberger, C. (2006) The structural basis of sirtuin substrate affinity. *Biochemistry* 45, 7511–7521.
- (38) Hawse, W. F., Hoff, K. G., Fatkins, D. G., Daines, A., Zubkova, O. V., Schramm, V. L., Zheng, W., and Wolberger, C. (2008) Structural insights into intermediate steps in the Sir2 deacetylation reaction. *Structure* 16, 1368–1377.
- (39) Hoff, K. G., Avalos, J. L., Sens, K., and Wolberger, C. (2006) Insights into the sirtuin mechanism from ternary complexes containing NAD⁺ and acetylated peptide. *Structure* 14, 1231–1240.
- (40) Sanders, B. D., Zhao, K., Slama, J. T., and Marmorstein, R. (2007) Structural basis for nicotinamide inhibition and base exchange in Sir2 enzymes. *Mol. Cell* 25, 463–472.
- (41) Zhao, K., Chai, X., and Marmorstein, R. (2003) Structure of the yeast Hst2 protein deacetylase in ternary complex with 2'-O-acetyl ADP ribose and histone peptide. *Structure (Cambridge)* 11, 1403–1411.
- (42) Zhao, K., Chai, X., and Marmorstein, R. (2004) Structure and substrate binding properties of cobB, a Sir2 homolog protein deacetylase from *Escherichia coli*. *J. Mol. Biol.* 337, 731–741.
- (43) Zhao, K., Harshaw, R., Chai, X., and Marmorstein, R. (2004) Structural basis for nicotinamide cleavage and ADP-ribose transfer by NAD⁺-dependent Sir2 histone/protein deacetylases. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8563–8568.
- (44) Schwartz, D., Chou, M. F., and Church, G. M. (2009) Predicting protein post-translational modifications using meta-analysis of proteome scale data sets. *Mol. Cell Proteomics* 8, 365–379.
- (45) Cherkasov, A., Hilpert, K., Jenssen, H., Fjell, C. D., Waldbrook, M., Mullaly, S. C., Volkmer, R., and Hancock, R. E. (2009) Use of artificial intelligence in the design of small peptide antibiotics effective against a broad spectrum of highly antibiotic-resistant superbugs. *ACS Chem. Biol.* 4, 65–74.
- (46) Fjell, C. D., Jenssen, H., Hilpert, K., Cheung, W. A., Pante, N., Hancock, R. E., and Cherkasov, A. (2009) Identification of novel antibacterial peptides by chemoinformatics and machine learning. *J. Med. Chem.* 52, 2006–2015.
- (47) Bao, J., Lu, Z., Joseph, J. J., Carabenciov, D., Dimond, C. C., Pang, L., Samsel, L., McCoy, J. P., Leclerc, J., Nguyen, P., Gius, D., and Sack, M. N. (20XX) Characterization of the murine SIRT3 mitochondrial localization sequence and comparison of mitochondrial enrichment and deacetylase activity of long and short SIRT3 isoforms. *J. Cell. Biochem.* 110, 238–247.
- (48) Cooper, H. M., and Spelbrink, J. N. (2008) The human SIRT3 protein deacetylase is exclusively mitochondrial. *Biochem. J.* 411, 279–285.
- (49) Onyango, P., Celic, I., McCaffery, J. M., Boeke, J. D., and Feinberg, A. P. (2002) SIRT3, a human SIR2 homologue, is an NAD-dependent deacetylase localized to mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13653–13658.
- (50) Schwer, B., North, B. J., Frye, R. A., Ott, M., and Verdin, E. (2002) The human silent information regulator (Sir)2 homologue hSIRT3 is a mitochondrial nicotinamide adenine dinucleotide-dependent deacetylase. *J. Cell Biol.* 158, 647–657.
- (51) Sweet, R. M., and Eisenberg, D. (1983) Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure. *J. Mol. Biol.* 171, 479–488.
- (52) Bhaskaran, R., and Ponnuswamy, P. K. (1988) Positional flexibilities of amino acid residues in globular proteins. *Int. J. Pept. Protein Res.* 32, 241–255.
- (53) Deleage, G., and Roux, B. (1987) An algorithm for protein secondary structure prediction based on class prediction. *Protein Eng.* 1, 289–294.
- (54) Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H., and Zehfus, M. H. (1985) Hydrophobicity of amino acid residues in globular proteins. *Science* 229, 834–838.
- (55) Janin, J. (1979) Surface and inside volumes in globular proteins. *Nature* 277, 491–492.
- (56) Lifson, S., and Sander, C. (1979) Antiparallel and parallel β -strands differ in amino acid residue preferences. *Nature* 282, 109–111.
- (57) Liu, D. C., and Nocedal, J. (1989) On the limited memory BFGS method for large scale optimization methods. *Math. Programming* 45, 503–528.
- (58) Mitchell, T. M. (1997) *Machine Learning*, McGraw-Hill, New York.
- (59) Tibshirani, R. (1996) Regression shrinkage and selection via the lasso. *J. R. Stat. Soc., Ser. B* 58, 267–288.

(60) Smith, B. C., Hallows, W. C., and Denu, J. M. (2009) A continuous microplate assay for sirtuins and nicotinamide-producing enzymes. *Anal. Biochem.* 394, 101–109.

(61) Weiser, A. A., Or-Guil, M., Tapia, V., Leichsenring, A., Schuchhardt, J., Frommel, C., and Volkmer-Engert, R. (2005) SPOT synthesis: reliability of array-based measurement of peptide binding affinity. *Anal. Biochem.* 342, 300–311.

(62) Tapia, V., Bongartz, J., Schutkowski, M., Bruni, N., Weiser, A., Ay, B., Volkmer, R., and Or-Guil, M. (2007) Affinity profiling using the peptide microarray technology: a case study. *Anal. Biochem.* 363, 108–118.

(63) Jin, L., Wei, W., Jiang, Y., Peng, H., Cai, J., Mao, C., Dai, H., Choy, W., Bemis, J. E., Jirousek, M. R., Milne, J. C., Westphal, C. H., and Perni, R. B. (2009) Crystal structures of human SIRT3 displaying substrate-induced conformational changes. *J. Biol. Chem.* 284, 24394–24405.

(64) Khan, A. N., and Lewis, P. N. (2005) Unstructured conformations are a substrate requirement for the Sir2 family of NAD-dependent protein deacetylases. *J. Biol. Chem.* 280, 36073–36078.

(65) Lins, L., and Brasseur, R. (1995) The hydrophobic effect in protein folding. *FASEB J.* 9, 535–540.

(66) Kouzarides, T. (2000) Acetylation: a regulatory modification to rival phosphorylation?. *EMBO J.* 19, 1176–1179.

(67) Yu, W., Lin, Y., Yao, J., Huang, W., Lei, Q., Xiong, Y., Zhao, S., and Guan, K. L. (2009) Lysine 88 acetylation negatively regulates ornithine carbamoyltransferase activity in response to nutrient signals. *J. Biol. Chem.* 284, 13669–13675.

(68) Law, I. K., Liu, L., Xu, A., Lam, K. S., Vanhoutte, P. M., Che, C. M., Leung, P. T., and Wang, Y. (2009) Identification and characterization of proteins interacting with SIRT1 and SIRT3: implications in the anti-aging and metabolic effects of sirtuins. *Proteomics* 9, 2444–2456.

(69) Jackson, M. D., and Denu, J. M. (2002) Structural identification of 2'- and 3'-O-acetyl-ADP-ribose as novel metabolites derived from the Sir2 family of β -NAD⁺-dependent histone/protein deacetylases. *J. Biol. Chem.* 277, 18535–18544.

(70) Jackson, M. D., Schmidt, M. T., Oppenheimer, N. J., and Denu, J. M. (2003) Mechanism of nicotinamide inhibition and transglycosidation by Sir2 histone/protein deacetylases. *J. Biol. Chem.* 278, 50985–50998.

(71) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254.

(72) Nikolovska-Coleska, Z., Wang, R., Fang, X., Pan, H., Tomita, Y., Li, P., Roller, P. P., Krajewski, K., Saito, N. G., Stuckey, J. A., and Wang, S. (2004) Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. *Anal. Biochem.* 332, 261–273.

(73) Roehrl, M. H., Wang, J. Y., and Wagner, G. (2004) A general framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization. *Biochemistry* 43, 16056–16066.

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

(75) DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.

(76) Miller, B. T., Singh, R. P., Klauda, J. B., Hodoscek, M., Brooks, B. R., and Woodcock, H. L. (2008) CHARMMing: a new, flexible web portal for CHARMM. *J. Chem. Inf. Model.* 48, 1920–1929.