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

SUPPORTING INFORMATION

Contents:

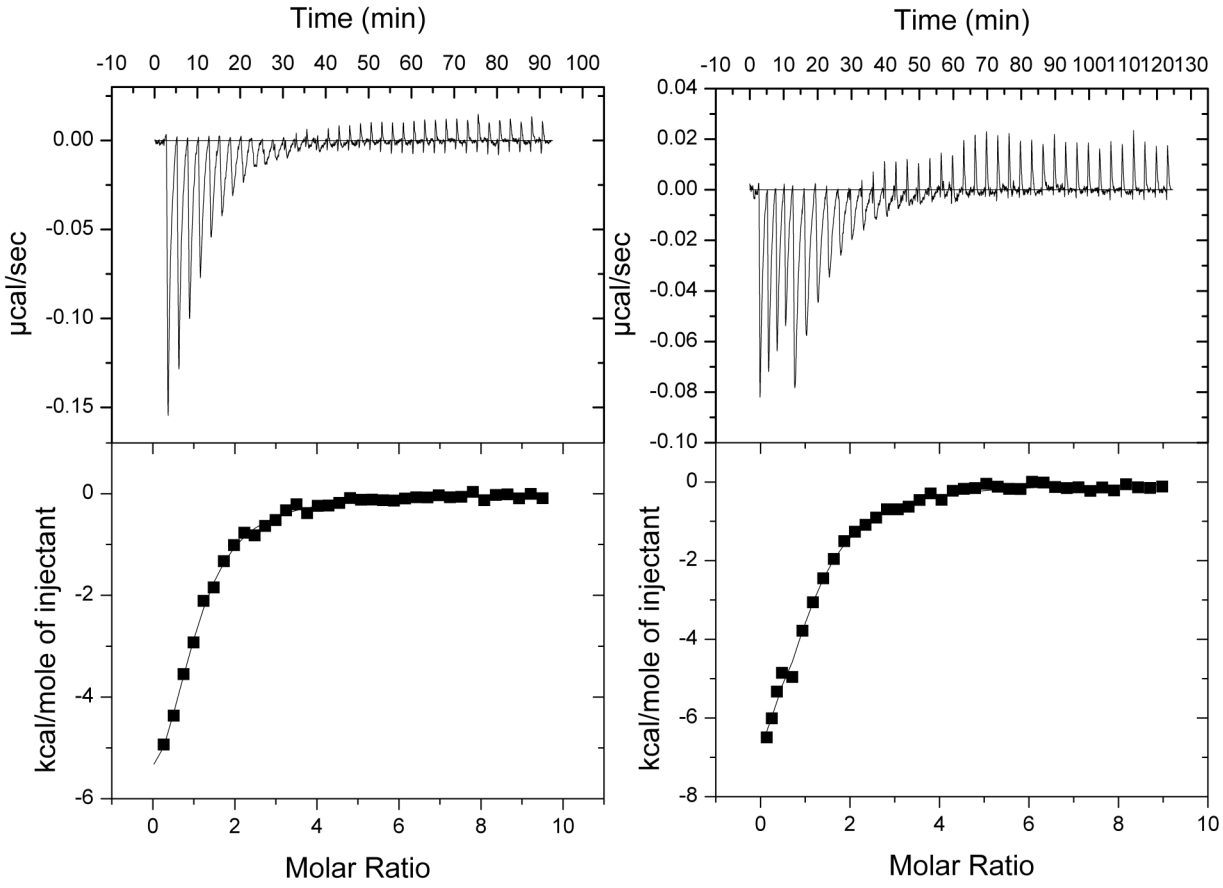


Figure S1. Representative isothermal titration calorimetry curve for thiotrifluoroacetyl- and pentafluoropropionyl-lysine peptide binding to Hst2. A solution of thiotrifluoroacetyl-lysine (left) or pentafluoropropionyl-lysine (right) H3 11-mer peptide was successively injected into the cell containing Hst2 and heats of binding were measured after each injection. The least-squared fit to the binding parameters ΔH° and K_d was determined using Origin software.

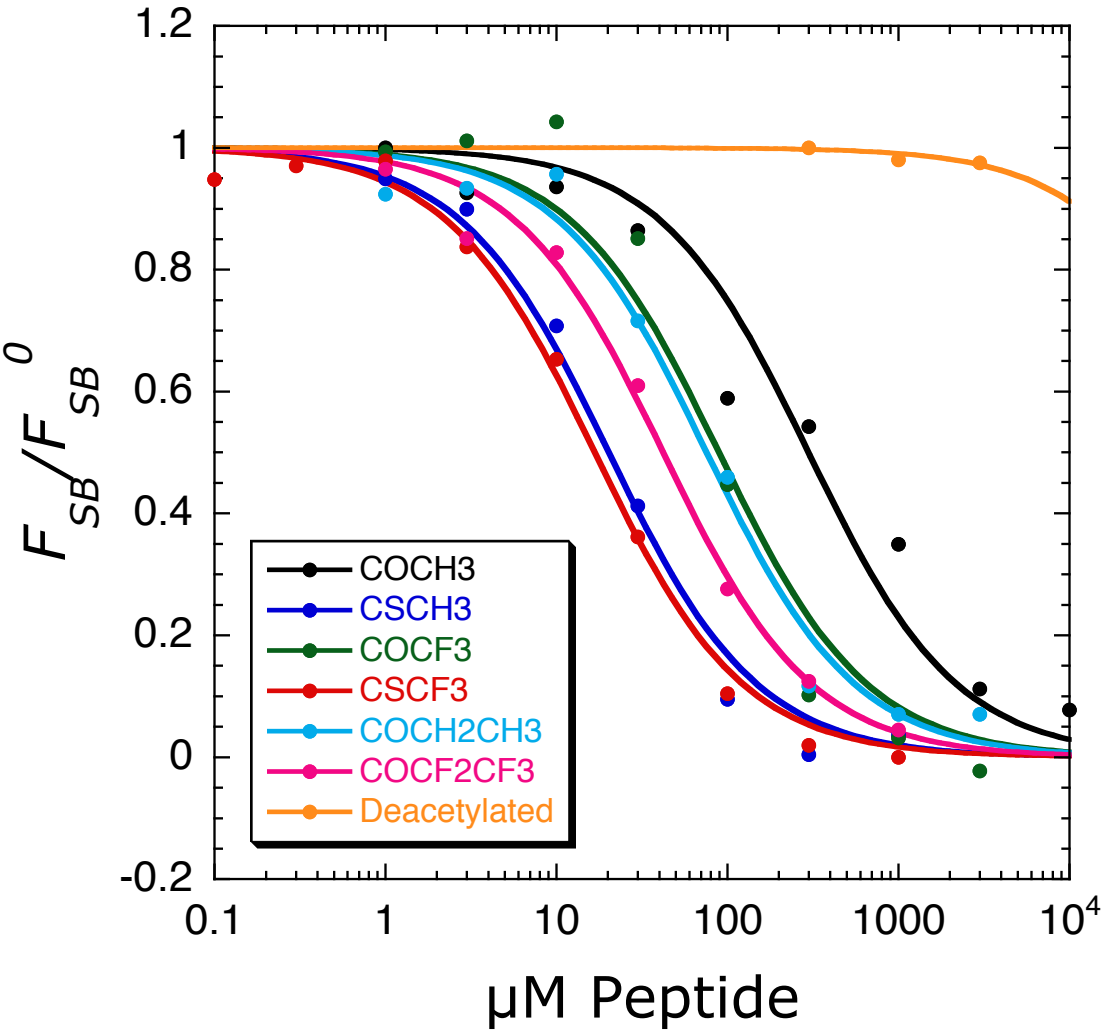


Figure S2. Competitive binding curves for various acetyl-lysine analog peptides. F_{SB} is the fraction of Fluor-ACS2 bound to SIRT3 at a given competitor peptide concentration and F_{SB}^{0} is the fraction of Fluor-ACS bound to SIRT3 at zero competitor peptide concentration. SIRT3 (10 μM) was incubated with 20 nM Fluor-ACS2 peptide and various concentrations unlabeled acetyl-lysine analog peptides (H_2N-KSTGGK(acetyl analog)APRKQ-OH) and the fluorescence anisotropy determined. Dissociation constants were then calculated as detailed under Methods.

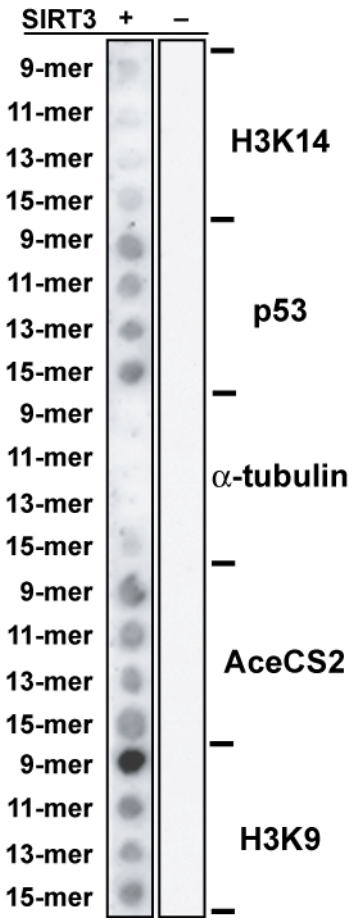


Figure S3. Comparison of peptide length for SIRT3 binding to SPOT libraries. Peptide SPOT libraries of various known sirtuin substrates and varied peptide length and a central thiotrifluoroacetyl-lysine residue were probed with and without SIRT3 as detailed under Experimental Procedures. The 9-mer peptide shows greatest affinity with the least nonspecific binding for SIRT3.

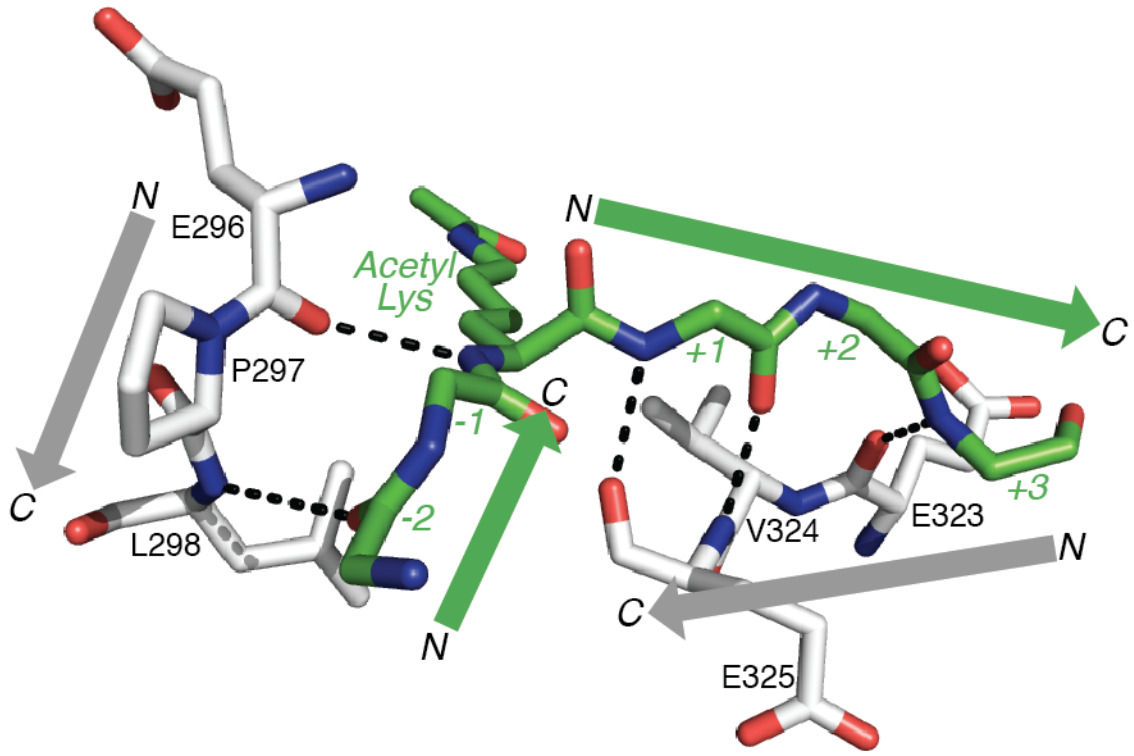


Figure S4. The acetyl-lysine substrate peptide binds to sirtuins in an antiparallel beta strand fashion. The acetylated acetyl-CoA synthetase 2 peptide bound to SIRT3 (PDB code 3GLR; (*1*)) is shown in green with sidechains removed for clarity. The interacting residues from SIRT3 are shown in white. Putative hydrogen bonds are shown as black dashed lines. Images were generated using Pymol (*2*).

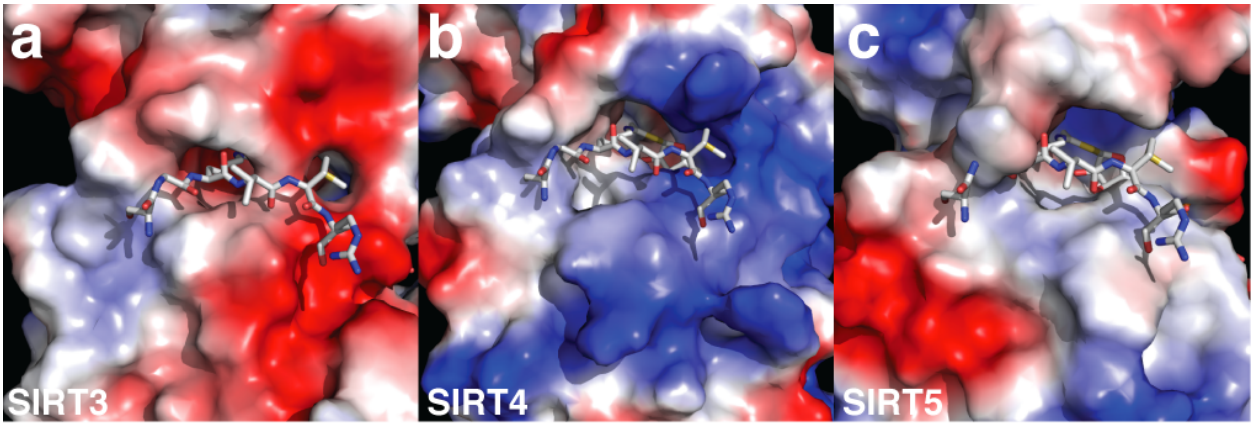
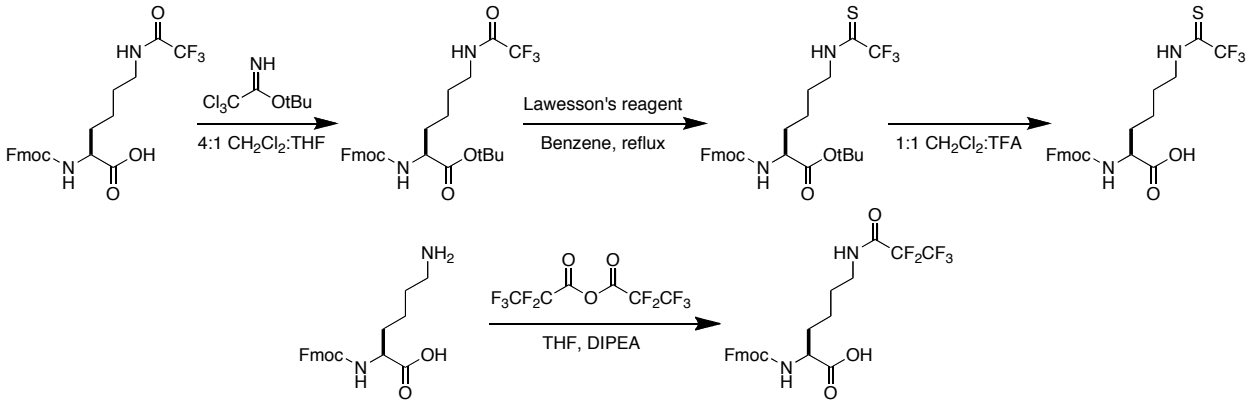


Figure S5. Comparison of SIRT3, SIRT4, and SIRT5 substrate binding sites. (a) SIRT3 was modeled from a structure of a thioalkylamidate bound to SIRT3 resulting from the reaction of a thioacetyl-lysine acetyl-CoA synthetase 2 (ACS2) peptide and NAD^+ (PDB code 3GLT) (*1*). (b) SIRT4 is shown as a homology model based on the template structure of a Sir2 homolog from *Archaeoglobus fulgidus* (PDB code 1S7G) (*3*) using SWISS-MODEL (*4, 5*). (c) SIRT5 was modeled from a structure of SIRT5 with ADP-ribose bound within the active-site (PDB code 2B4Y) (*6*). As the SIRT4 homology model and SIRT5 structure do not have a peptide substrate bound, the thioalkylamidate-ACS2 peptide was modeled from the SIRT3 structure to identify the approximate peptide-binding site. The enzyme surface is colored based on electrostatic potential with red, blue, and white surface representing negative, positive, and hydrophobic residues, respectively. Images were generated using Pymol (*2*).

Table S1. Mass spectral characterization and k_{cat}/K_m values for *in vitro* validation peptides.

Sequence	**Mass expected** $[M+H]^+$ (m/z)	**Mass observed** $[M+H]^+$ (m/z)	k_{cat}/K_m **(M^{-1} s^{-1})**
IINQK(ac)RFND	1188.6	1188.8	10093
RFNWK(ac)LFWQ	1365.7	1365.9	10045
TRSGK(ac)VMRR	1131.6	1131.4	8241
PLSRK(ac)HGGP	989.6	989.8	7470
RRYIK(ac)WPLL	1285.8	1286.0	6221
SLERK(ac)FGKH	1142.6	1143.1	5056
EIFDK(ac)HYKT	1221.6	1221.7	4824
RQVVK(ac)KYWA	1218.7	1218.6	4661
FLWIK(ac)VKGI	1144.7	1145.3	4626
VDFSK(ac)ALGK	1005.6	1005.5	4327
TYIDK(ac)WFLY	1289.6	1290.3	3934
RTAMK(ac)YNLG	1094.6	1095.2	3052
LATQK(ac)YSVA	1021.6	1022.2	1503
DAKSK(ac)EFAQ	1064.5	1064.6	1250
DVFTK(ac)GYGF	1074.5	1074.5	1164
EQYAK(ac)ENGT	1080.5	1080.9	869
TEEEK(ac)NIKW	1217.6	1218.3	427
SPGPK(ac)DEDE	1014.4	1014.6	410
VGEDK(ac)YGNK	1050.5	1051.0	255
EEREK(ac)KPAE	1156.6	1156.4	235
EEEEK(ac)FDPN	1177.5	1177.7	139
EESDK(ac)RESE	1149.5	1149.8	44
EEKEK(ac)AKEK	1159.6	1159.8	37
KNEEK(ac)EEDD	1176.5	1176.7	36

Scheme S1. Synthesis of Fmoc-Lys($CSCF_3$)-OH and Fmoc-Lys($COCF_2CF_3$)-OH



Supplemental Methods

Fmoc-Lys($COCF_3$)-OtBu. Fmoc-Lys($COCF_3$)-OH (7.89 g, 17.0 mmol) was dissolved in 56.6 mL of 4:1 by volume anhydrous dichloromethane (CH_2Cl_2)/tetrahydrofuran (THF). To the stirring solution was added *tert*-butyl 2,2,2-trichloroacetimidate (6.08 mL, 34.0 mmol, 2 equiv). The reaction was stirred overnight at room temperature under argon. The reaction was then concentrated under reduced pressure and redissolved in ethyl acetate (EtOAc) (200 mL). The solution was washed with 1:1 H_2O:saturated sodium bicarbonate ($NaHCO_3$) (200 mL) and then brine (200 mL). The organic layer was dried over anhydrous $MgSO_4$, filtered, and concentrated. The crude product was purified by silica gel flash chromatography (3:1 hexanes/EtOAc) to give Fmoc-Lys($COCF_3$)-OtBu in quantitative yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ ppm: 7.75 (d, J_{HH} = 7.5 Hz, 2H), 7.59 (d, J_{HH} = 7.5 Hz, 2H), 7.39 (t, J_{HH} = 7.5 Hz, 2H), 7.30 (t, J_{HH} = 7.5 Hz, 2H), 6.73 (bs, 1H), 6.61 (bs, 1H), 5.53 (d, J_{HH} = 8 Hz, 1H), 4.38 (m, 2H), 4.23 (m, 2H), 3.34 (m, 2H), 1.84–1.40 (m, 6H), 1.47 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm: 171.6, 157.4 (q, J_{CF} = 36.6 Hz), 156.2, 143.8, 141.3, 127.7, 127.1, 125.1, 120.0, 115.9 (q, J_{CF} = 286 Hz), 91.9, 82.5, 67.0, 53.9, 47.1, 39.6, 32.5, 28.0, 22.2.

Fmoc-Lys($CSCF_3$)-OtBu. Fmoc-Lys($COCF_3$)-OtBu (2.38 g, 4.6 mmol) was dissolved in 45.7 mL of anhydrous benzene. To the stirring solution was added Lawesson's reagent (2.03 g, 5.0 mmol, 1.1 equiv). The reaction was refluxed for 4 hours. The reaction was then allowed to cool to room temperature and concentrated under reduced pressure. The crude product was purified on a short plug of silica gel (4:1 hexanes:EtOAc). Fractions containing product were concentrated under reduced pressure and further purified by silica gel flash chromatography (17:3 hexanes/EtOAc) to give Fmoc-Lys($CSCF_3$)-OtBu (2.38 g) in 97% yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ ppm: 7.76 (d, J_{HH} = 7.5 Hz, 2H), 7.59 (d, J_{HH} = 7.5 Hz, 2H), 7.40 (t, J_{HH} = 7.5 Hz, 2H), 7.31 (t, J_{HH} = 7.5 Hz, 2H), 6.68 (bs, 1H), 6.45 (bs, 1H), 5.52 (d, J_{HH} = 8 Hz, 1H), 4.40 (m, 2H), 4.25 (m, 2H), 3.65 (m, 2H), 1.88–1.45 (m, 6H), 1.48 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm: 183.3 (q, J_{CF} = 35.5 Hz), 171.5, 156.4, 143.8, 141.3, 127.8, 127.1, 125.0, 120.0, 117.4 (q, J_{CF} = 278 Hz), 82.6, 67.0, 53.6, 47.1, 45.7, 32.8, 28.0, 26.2, 22.5.

Fmoc-Lys($CSCF_3$)-OH. Fmoc-Lys($CSCF_3$)-OtBu (5.2 g, 9.7 mmol) was dissolved in 1:1 by volume trifluoroacetic acid (TFA): CH_2Cl_2 (64.6 mL). The reaction was stirred 4 hours at room temperature. The reaction was then concentrated under reduced pressure. The crude product was purified by silica gel flash chromatography (98:1:1 CH_2Cl_2:methanol:acetic acid) to give Fmoc-Lys($COCF_3$)-OtBu as a pale yellow foam in quantitative yield. ^{1}H NMR (500 MHz, DMSO-d_6) δ ppm: 12.56 (bs, 1H), 11.23 (bs, 1H), 7.88 (d, J_{HH} = 7.5 Hz, 2H), 7.72 (dd, J_{HH} = 7.5, 2.5 Hz, 2H), 7.62 (d, J_{HH} = 8 Hz, 1H), 7.41 (t, J_{HH} = 7.5 Hz, 2H), 7.32 (td, J_{HH} = 7.5, 1 Hz, 2H), 4.28 (d, J_{HH} = 7 Hz, 2H), 4.22 (t, J_{HH} = 7 Hz, 1H), 3.93 (td, J_{HH} = 8, 4.5 Hz, 1H), 3.58 (m, 2H), 1.90–1.32 (m, 6H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm: 183.5 (q, J_{CF} = 35 Hz), 176.6, 156.5, 143.6, 141.3, 127.8, 127.1, 125.0, 120.1, 117.4 (q, J_{CF} = 278 Hz), 67.3, 53.1, 47.0, 45.5, 32.2, 26.2, 22.6. ESI- MS: $[C_{23}H_{22}F_3N_2O_4S]^-$ 479.1428 (obsd), 479.1258 (calcd).

Fmoc-Lys($COCF_2CF_3$)-OH. Fmoc-Lys-OH (400 mg, 0.99 mmol, 1 equiv) was dissolved in THF (9.9 mL) and cooled to 0 ºC in an ice bath. DIPEA (344 µL, 1.98 mmol, 2 equiv) and pentafluoropropionic anhydride (195 µL, 0.99 mmol, 1 equiv) were added slowly and the reaction mixture was stirred for 0.5 hr. The reaction was then concentrated under reduced pressure and redissolved in ethyl acetate (50 mL). The solution was washed with 1 M HCl (50 mL) and brine (50 mL). The organic layer was dried over anhydrous $MgSO_4$, filtered, and concentrated. The crude product was purified by silica gel flash chromatography (0 to 2% v/v methanol in 99:1 CH_2Cl_2:acetic acid) to give Fmoc-Lys($COCF_2CF_3$)-OH (86 mg) as a white powder in 16% yield. ^{1}H NMR (500 MHz, CD_3OD) δ ppm: 7.74 (d, J_{HH} = 8 Hz, 2H), 7.63 (t, J_{HH} = 7.5, 2H), 7.35 (t, J_{HH} = 7.5 Hz, 2H), 7.27 (t, J_{HH} = 7.5, 2H), 4.31 (m, 2H), 4.15 (m, 2H), 1.88–1.37 (m, 8H). ^{13}C NMR (125 MHz, CD_3OD) δ ppm: 174.5, 158.0 (t, J_{CF} = 25 Hz), 157.3, 143.8, 141.2, 127.4, 126.8, 124.8, 119.5, 118.0 (qt, J_{CF} = 284, 38 Hz), 107.0 (tq, J_{CF} = 284, 38 Hz), 66.6, 53.8, 47.0, 39.2, 30.8, 28.0, 22.7. ESI-MS: $[C_{24}H_{22}F_5N_2O_5]^-$ 513.1443 (obsd), 513.1454 (calcd).

Solid-phase peptide synthesis of Fluor-ACS2 peptide. Fluor-ACS2 peptide (Fluorescein-LPKTRSGK($CSCF_3$)VMRR-OH) was synthesized using standard tBu/Fmoc solid-phase peptide synthesis (SPPS) techniques by the University of Wisconsin-Madison Peptide Synthesis Facility. Each amino acid was coupled using ~5 equivalents activated amino acid (1 equiv HBTU and 2 equiv NMM). The thiotrifluoroacetyl residue was incorporated into the peptide with Fmoc-Lys($CSCF_3$)-OH. 5(6)-carboxyfluorescein was coupled to the N-terminus of the peptide using HBTU/NMM coupling conditions in ~3-fold molar excess. After completion of the synthesis, the resin was rinsed with dichloromethane and dried. The full-length peptide was then cleaved from the resin and deprotected with 92.5% TFA, 2.5% ethanethiol, and 5% thioanisole. The peptide was precipitated with cold t-butylmethylether and sedimented by centrifuge washing twice with additional t-butylmethylether. The precipitate was dried, redissolved in water, and lyophilized. Mass of the cleaved peptide was confirmed by MALDI-TOF mass spectrometry. Crude peptides were purified by semipreparative RP- HPLC on a C18 small pore column (Grace Vydac, 10 × 250 mm, 10 μm) eluting with a gradient of 0-100% acetonitrile with 0.02% TFA (solvent B) in water with 0.05% TFA (solvent A). Fractions collected were lyophilized to yield Fluor-ACS2 as a dry bright yellow powder. Fluor-ACS2 peptide concentrations were determined by fluorescein absorbance at 494 nm with an extinction coefficient of 68,000 M^{-1} cm^{-1}.

Solid-phase peptide synthesis of acetyl-lysine analog peptides. Synthetic acetyl-lysine analog H3 peptides (H_2N-KSTGGK(acetyl analog)APRKQ-OH) were synthesized using standard tBu/Fmoc solid-phase peptide synthesis (SPPS) techniques by the University of Wisconsin-Madison Peptide Synthesis Facility) as previously described (*7*). The thiotrifluoroacetyl and pentafluoropropionyl derivatives were synthesized from Fmoc-Lys($CSCF_3$)-OH and Fmoc-Lys($COCF_2CF_3$)-OH. After completion of the synthesis, the resin was rinsed with dichloromethane and dried. The full-length peptide was then cleaved from the resin and deprotected with 92.5% TFA, 2.5% ethanethiol, and 5% thioanisole. The peptide was precipitated with cold t-butylmethylether and sedimented by centrifuge washing twice with additional t-butylmethylether. The precipitate was dried, redissolved in water, and lyophilized. Masses of the cleaved peptides were confirmed by MALDI-TOF mass spectrometry. Thiotrifluoroacetyl H3 peptide MALDI-MS: $[C_{50}H_{88}F_3N_{18}O_{15}S]^+$ 1269.4 (obsd), 1269.6 (calcd). Pentafluoropropionyl H3 peptide MALDI-MS: $[C_{51}H_{88}F_5N_{18}O_{16}]^+$ 1303.3 (obsd), 1303.7 (calcd).

Crude peptides were purified by semipreparative RP- HPLC on a C18 small pore column (Grace Vydac, 10 × 250 mm, 10 μm) eluting with a gradient of 0-40% acetonitrile with 0.02% TFA (solvent B) in water with 0.05% TFA (solvent A). Fractions collected were lyophilized to yield final peptides as dry white powders of >85% purity. Peptide concentrations were determined from the weight of the peptide TFA salt assuming 1 equiv of TFA per cationic group.

Direct binding fluorescence polarization of Fluor-ACS2 to SIRT3. Fluor-ACS2 peptide (20 nM) was incubated with various SIRT3 concentrations from 100 nM to 800 μM in 50 mM Tris-HCl (pH 7.5) and 10% (v/v) glycerol in 100 μL volumes. Fluorescence polarization was measured at 25 ˚C using a Panvera Beacon 2000 FP system with 494??? nm excitation and 521??? nm emission wavelengths. Fluorescence anisotropy values (mA) were converted to fraction Fluor-ACS2 bound (F_{SB}) using the following equation previously described (*8*):

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

where A_{OBS} is the measured anisotrophy at the particular competitor peptide concentration, A_F is the anisotrophy of free Fluor-ACS2, A_B is the anisotrophy of the fully SIRT3 bound Fluor-ACS2, and Q is the ratio of fluorescence intensities of bound and free Fluor-ACS2. F_{SB} was then plotted versus SIRT3

concentration and fitted using the following equation to determine the dissociation constant (K_d) as previously described (*8*):

$$F_{SB} = \frac{K_d + L_{ST} + R_T - \sqrt{(K_d + L_{ST} + R_T)^2 - 4L_{ST}R_T}}{2L_{ST}}$$

where L_{ST} is the total Fluor-ACS2 concentration (20 nM) and R_T is the concentration of SIRT3.

Isothermal Titration Calorimetry of Acetyl-lysine Analog Peptide Binding to Hst2. The K_d value for 11-mer thiotrifluoroacetyl- and pentafluoropropionyl-lysine peptides based on the sequence of histone H3 acylated at lysine-14 (NH_2-KSTGGK($CSCF_3$)APRKQ-OH) binding to Hst2 were determined as described previously (*7, 9, 10*) using a MicroCal VP-ITC instrument.

Supplemental References

1. 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.
2. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA, USA.
3. Avalos, J. L., Boeke, J. D., and Wolberger, C. (2004) Structural basis for the mechanism and regulation of Sir2 enzymes, *Mol Cell 13*, 639-648.
4. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, *Bioinformatics 22*, 195-201.
5. Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., and Schwede, T. (2009) The SWISS-MODEL Repository and associated resources, *Nucleic Acids Res 37*, D387-392.
6. Schuetz, A., Min, J., Antoshenko, T., Wang, C. L., Allali-Hassani, A., Dong, A., Loppnau, P., Vedadi, M., Bochkarev, A., Sternglanz, R., and Plotnikov, A. N. (2007) Structural basis of inhibition of the human NAD+-dependent deacetylase SIRT5 by suramin, *Structure 15*, 377-389.
7. 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.
8. 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.
9. Smith, B. C., and Denu, J. M. (2007) Mechanism-based inhibition of sir2 deacetylases by thioacetyl-lysine Peptide, *Biochemistry 46*, 14478-14486.
10. Smith, B. C., and Denu, J. M. (2007) Acetyl-lysine Analog Peptides as Mechanistic Probes of Protein Deacetylases, *J Biol Chem 282*, 37256-37265.