--- SUPPLEMENTARY INFORMATION ---

# Histone deacetylase 6 structure and molecular basis of catalysis and inhibition

Yang Hai<sup>1</sup> & David W. Christianson<sup>1,2\*</sup>

<sup>1</sup>Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323, USA. <sup>2</sup>Radcliffe Institute for Advanced Study and Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA. \*e-mail: chris@sas.upenn.edu

### SUPPLEMENTARY RESULTS

Substrate 1	K <sub>м</sub> (μM)	$k_{\rm cat}$ (s <sup>-1</sup> )	<i>k</i> <sub>cat</sub> /K <sub>M</sub> (M <sup>-1</sup> s <sup>-1</sup> )
hCD12	22 ± 6	0.05 ± 0.01	$(2.3 \pm 0.9) \times 10^3$
hCD12 Y386F	26 ± 9	$0.06 \pm 0.02$	(2.3 ± 0.9) x 10 <sup>3</sup>
hCD12 Y782F	62 ± 4	$0.00030 \pm 0.00002$	$4.8 \pm 0.3$
hCD12 Y386F/Y782F	69 ± 3	$0.00021 \pm 0.00001$	$3.0 \pm 0.2$
zCD12	38 ± 6	0.8 ± 0.1	(2.10 ± 0.08) x 10 <sup>4</sup>
zCD12 H194A	31 ± 9	0.7 ± 0.1	$(2.3 \pm 0.8) \times 10^4$
zCD12 H574A	65 ± 9	$0.082 \pm 0.007$	$(1.2 \pm 0.1) \times 10^3$
zCD1	87 ± 9	$0.079 \pm 0.006$	$(0.9 \pm 0.1) \times 10^3$
zCD2	29 ± 9	$0.56 \pm 0.09$	$(1.9 \pm 0.5) \times 10^4$
Substrate 8	K <sub>Μ</sub> (μM)	<i>k</i> <sub>cat</sub> (s <sup>-1</sup> )	<i>k</i> <sub>cat</sub> /K <sub>M</sub> (M⁻¹s⁻¹)
hFLWT <sup>b</sup>	7 ± 4	0.04 ± 0.01	$(6 \pm 1) \times 10^3$
hCD12	11 ± 4	$0.049 \pm 0.009$	$(4.4 \pm 0.9) \times 10^3$
hCD1	n.a.	n.a.	n.a.
hCD2	22 ± 3	$0.020 \pm 0.002$	(0.91 ± 0.03) x 10 <sup>3</sup>
hCD12 Y386F	16 ± 4	$0.06 \pm 0.02$	$(4 \pm 1) \times 10^3$
hCD12 Y782F	42 ± 4	$0.0003 \pm 0.0001$	7 ± 2
hCD12 Y386F/Y782F	39 ± 3	$0.0004 \pm 0.0002$	9 ± 2
hCD12 Y225F/Y782F	56 ± 9	$0.048 \pm 0.004$	(0.86 ± 0.06) x 10 <sup>3</sup>
hCD12 K353L/Y782F	52 ± 6	$0.002 \pm 0.001$	40 ± 10
hCD2 D567A	26 ± 3	$0.022 \pm 0.006$	$(0.8 \pm 0.1) \times 10^3$
hCD2 S568A	21 ± 5	$0.0020 \pm 0.0007$	90 ± 10
zCD12	28 ± 3	$0.98 \pm 0.09$	$(3.5 \pm 0.5) \times 10^4$
zCD12 H194A	19 ± 6	$0.60 \pm 0.05$	$(3.2 \pm 0.4) \times 10^4$
zCD12 H574A	38 ± 7	$0.23 \pm 0.07$	$(6.2 \pm 0.9) \times 10^3$
zCD1	43 ± 6	$0.20 \pm 0.06$	$(4.7 \pm 0.3) \times 10^3$
zCD1 K330L	32 ± 7	$0.155 \pm 0.009$	$(5 \pm 1) \times 10^3$
zCD1 H82F/F202Y	47 ± 4	$0.016 \pm 0.006$	(0.34 ± 0.9) x 10 <sup>3</sup>
zCD2	22 ± 9	$0.69 \pm 0.08$	$(3.1 \pm 0.5) \times 10^4$
zCD2 H573A	8 ± 3	0.00021 ± 0.00004	26 ± 4
zCD2 H574A	9 ± 2	0.0006 ± 0.0001	67 ± 6
zCD2 N530A	40 ± 10	0.7 ± 0.1	$(1.7 \pm 0.9) \times 10^4$
zCD2 S531A	240 ± 40	0.03 ± 0.01	$(0.12 \pm 0.04) \times 10^3$

## Supplementary Table 1. Steady-state kinetics of HDAC6 catalytic domains<sup>a</sup>

Substrate 9	K <sub>Μ</sub> (μM)	$k_{\text{cat}}$ (s <sup>-1</sup> )	<i>k</i> <sub>cat</sub> /K <sub>M</sub> (M <sup>-1</sup> s <sup>-1</sup> )
hCD12	100 ± 20	0.26 ± 0.02	$(2.7 \pm 0.5) \times 10^3$
hCD12 Y386F	90 ± 40	0.18 ± 0.04	$(1.9 \pm 0.9) \times 10^3$
hCD12 Y782F	50 ± 10	0.080 ± 0.006	$(1.5 \pm 0.4) \times 10^3$
zCD1	100 ± 20	0.9 ± 0.1	$(8.6 \pm 0.9) \times 10^3$
zCD2	140 ± 50	1.2 ± 0.3	$(8.0 \pm 2.0) \times 10^3$
zCD12	110 ± 30	2.0 ± 0.3	$(1.9 \pm 0.4) \times 10^4$
Substrate 10	K <sub>M</sub> (μM)	k <sub>cat</sub> (s⁻¹)	<i>k</i> <sub>cat</sub> /K <sub>M</sub> (M⁻¹s⁻¹)
hCD12	80 ± 20	$0.20 \pm 0.02$	$(2.3 \pm 0.4) \times 10^3$
hCD12 Y386F	120 ± 20	0.13 ± 0.01	$(2.3 \pm 0.4) \times 10^3$
hCD12 Y782F	80 ± 10	0.078 ± 0.005	(1.0 ± 0.1) x 10 <sup>3</sup>
zCD1	120 ± 40	0.9 ± 0.2	(7.0 ± 2.0) x 10 <sup>3</sup>
zCD2	110 ± 20	$0.82 \pm 0.08$	(7.3 ± 0.7) x 10 <sup>3</sup>
zCD12	80 ± 20	1.5 ± 0.1	$(1.9 \pm 0.2) \times 10^4$
Substrate 11	K <sub>Μ</sub> (μM)	$k_{\text{cat}}$ (s <sup>-1</sup> )	<i>k</i> <sub>cat</sub> /K <sub>M</sub> (M <sup>-1</sup> s <sup>-1</sup> )
hCD12	210 ± 70	0.11 ± 0.04	(0.59 ± 0.08) x 10 <sup>3</sup>
hCD12 Y386F	140 ± 60	$0.10 \pm 0.02$	$(0.7 \pm 0.2) \times 10^3$
hCD12 Y782F	n.a.	n.a.	n.a.
zCD1	n.d. <sup>d</sup>	n.d.	11 ± 6
zCD2	110 ± 20	$0.44 \pm 0.04$	$(4.0 \pm 0.4) \times 10^3$
zCD12	150 ± 30	$0.62 \pm 0.08$	$(4.2 \pm 0.5) \times 10^3$
Substrate 12	K <sub>Μ</sub> (μΜ)	$k_{\text{cat}}$ (s <sup>-1</sup> )	<i>k</i> <sub>cat</sub> /K <sub>M</sub> (M <sup>-1</sup> s <sup>-1</sup> )
hCD12	70 ± 30	0.25 ± 0.04	$(4 \pm 1) \times 10^3$
hCD12 Y386F	80 ± 30	0.26 ± 0.04	$(3.1 \pm 0.9) \times 10^3$
hCD12 Y782F	n.a.	n.a.	n.a.
zCD1	90 ± 30	0.047 ± 0.006	$(0.6 \pm 0.1) \times 10^3$
zCD2	80 ± 20	0.9 ± 0.1	$(1.2 \pm 0.2) \times 10^4$
zCD12	70 ± 10	1.3 ± 0.1	$(1.8 \pm 0.3) \times 10^4$
Substrate 13	K <sub>Μ</sub> (μΜ)	<i>k</i> <sub>cat</sub> (s <sup>-1</sup> )	<i>k</i> <sub>cat</sub> /K <sub>M</sub> (M <sup>-1</sup> s <sup>-1</sup> )
hCD12	70 ± 20	0.07 ± 0.01	$(0.9 \pm 0.1) \times 10^3$
hCD12 Y386F	60 ± 7	0.060 ± 0.002	(0.97 ± 0.08) x 10 <sup>3</sup>
hCD12 Y782F	n.d.	n.d.	3 ± 1
zCD1	n.d.	n.d.	20 ± 10
zCD2	130 ± 50	0.39 ± 0.07	$(3.0 \pm 0.7) \times 10^3$
zCD12	200 ± 40	0.58 ± 0.06	$(2.9 \pm 0.4) \times 10^3$

<sup>a</sup>Data represent mean values ± s.e.m. (n=3). <sup>b</sup>hFLWT, human full-length wild-type HDAC6 from Enzo Life Sciences. <sup>c</sup>n.a., no activity. <sup>d</sup>n.d., not determined.

	zCD1-TSA	zCD2-TSA	MBP-hCD2-TSA	Unliganded zCD2
	complex	complex	complex	
Data collection				
Space group	<i>P</i> 2 <sub>1</sub>	P21212	P212121	<i>P</i> 1
Cell dimensions				
a, b, c (Å),	52.9, 123.7, 55.2	83.9, 94.4, 51.7	49.3, 149.0, 216.3	48.5, 55.4, 74.2
α, β, γ ( <sup>°</sup> )	90.0, 113.8, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	73.6, 89.8, 82.6
Resolution (Å) <sup>*</sup>	45.04 - 2.15	50.00 - 1.59	149.00 - 2.79	50.00 - 2.00
	(2.24 - 2.15)	(1.66 - 1.59)	(2.90 - 2.79)	(2.08 - 2.00)
R <sub>merge</sub>	0.178 (0.701)	0.061 (0.033)	0.204 (1.853)	0.162 (0.501)
<i>R</i> <sub>pim</sub>	0.122 (0.493)	0.050 (0.297)	0.141 (0.607)	0.162 (0.509)
Ι/ σΙ	10.6 (2.0)	23.7 (4.3)	12.9 (1.9)	6.4 (1.8)
Completeness (%)	99.0 (95.9)	100.0 (100.0)	100.0 (100.0)	97.0 (96.6)
Redundancy <sup>a</sup>	3.1 (2.8)	2.5 (2.4)	11.0 (11.3)	1.9 (1.9)
CC <sub>1/2</sub>	0.754 (0.536)	0.935 (0.865)	0.986 (0.585)	0.967 (0.662)
Refinement				
Resolution (Å)	45.04 - 2.15	44.01 - 1.59	87.54 - 2.79	15.91 - 2.00
	(2.23 - 2.15)	(1.65 - 1.59)	(2.89 - 2.79)	(2.07 - 2.00)
No. reflections	34834 (3335)	55924 (5519)	40788 (4012)	47957 (4586)
R <sub>work</sub> / R <sub>free</sub>	0.211/0.255	0.129/0.163	0.213/0.275	0.224/0.272
	(0.282/0.325)	(0.128/0.187)	(0.313/0.377)	(0.276/0.316)
No. atoms				
Protein	5558	2836	3260	5541
Ligand/ion	63	77	77	12
Water	164	376	355	305
<i>B</i> factors (Å <sup>2</sup> )				
Protein	32	13	55	17
Ligand/ion	34	32	62	25
Water	30	28	59	21
R.m.s. deviations				
bond lengths (Å)	0.006	0.010	0.004	0.006
bond angles (°)	0.9	1.2	0.8	1.0

#### Supplementary Table 2. Data collection and refinement statistics.

Each dataset was collected from a single crystal. \*Values in parentheses are for highest-resolution shell.

	H574A zCD2-	Y785F zCD2-	zCD2-HC toxin	zCD2-trifluoro-	zCD2-acetate
	substrate 8	substrate 1	complex	ketone inhibitor	complex
Data collection	complex	complex		complex	
	<b>P</b> 2.	P2.2.2	<b>D2.2.2</b>	<b>D</b> 2.2.2.	P2.2.2.
Coll dimonsions	<i>I Z</i> <sub>1</sub>	r z1z1z	1 21212	I Z1Z1Z1	<i>r</i> z <sub>1</sub> z <sub>1</sub> z <sub>1</sub>
	FF 0 94 0 97 0	92 2 04 7 51 6	92 5 04 1 51 5	746 024 062	75 1 01 9 06 2
a, b, c (A),	00.0.09.1.00.0	00.0.00.000.0	00.0, 00.0, 00.0	74.0, 92.4, 90.2	75.1, 91.0, 90.2
α, p, γ ( )	90.0, 98.1, 90.0	90.0, 90.0 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (A)	50.00 - 1.80	50.00 - 1.82	50.00 - 1.73	50.00 - 2.16	50.00 - 2.25
_	(1.87 - 1.80)	(1.88 - 1.82)	(1.79 - 1.73)	(2.24 - 2.16)	(2.33 - 2.25)
R <sub>merge</sub>	0.142 (0.713)	0.164 (0.655)	0.151 (0.442)	0.148 (0.334)	0.130 (0.356)
R <sub>pim</sub>	0.098 (0.509)	0.049 (0.344)	0.061 (0.273)	0.061 (0.178)	0.055 (0.212)
Ι/ σΙ	9.9 (1.8)	13.6 (1.1)	14.0 (1.6)	13.5 (3.0)	14.5 (2.8)
Completeness (%)	98.9 (99.9)	99.4 (94.1)	98.7 (90.6)	98.8 (90.3)	99.3 (93.9)
Redundancy <sup>a</sup>	3.0 (3.0)	11.0 (4.1)	6.6 (3.1)	6.6 (3.9)	6.3 (3.5)
CC <sub>1/2</sub>	0.975 (0.524)	0.958 (0.692)	0.976 (0.808)	0.977 (0.933)	0.969 (0.856)
Refinement					
Resolution (Å)	12.34 - 1.80	45.32 - 1.81	47.07 - 1.73	49.71 - 2.16	49.77 - 2.24
	(1.87 - 1.80)	(1.88 - 1.81)	(1.79 - 1.73)	(2.24 - 2.16)	(2.33 - 2.24)
No. reflections	70922 (6942)	37343 (3333)	42831 (3774)	35829 (3223)	32149 (2879)
R <sub>work</sub> / R <sub>free</sub>	0.162/0.194	0.157/0.183	0.152/0.173	0.196/0.242	0.212/0.234
	(0.238/0.270)	(0.260/0.265)	(0.224/0.244)	(0.221/0.275)	(0.280/0.319)
No. atoms					
Protein	5718	2792	2817	5581	5603
Ligand/ion	28	33	25	24	22
Water	25	29	27	23	26
<i>B</i> factors (Å <sup>2</sup> )					
Protein	5718	2792	2817	5581	5603
Ligand/ion	28	33	25	24	22
Water	25	29	27	23	26
R.m.s. deviations					
bond lengths (Å)	0.008	0.009	0.012	0.005	0.005
bond angles (°)	1.0	1.1	1.2	1.0	1.0

#### Supplementary Table 3. Data collection and refinement statistics.

Each dataset was collected from a single crystal. \*Values in parentheses are for highest-resolution shell.

	zCD2-SAHA complex	zCD2-Belinostat complex	zCD2-HPOB complex	zCD2- Panobinostat complex	zCD2- Oxamflatin complex
Data collection					
Space group	<i>P</i> 2 <sub>1</sub>	P212121	<i>P</i> 1	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2	P21212
Cell dimensions					
a, b, c (Å),	54.8, 83.6, 86.8	76.0, 95.5, 96.3	48.7, 56.6, 74.8	65.2, 91.9, 140.4	83.5, 94.4, 51.6
α, β, γ ( <sup>°</sup> )	90.0, 98.1, 90.0	90.0, 90.0, 90.0	106.4, 90.1, 97.1	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution $(Å)^{*}$	50.00 - 1.32	50.00 - 1.85	50.60 - 1.90	50.00 - 2.60	50.00 - 2.54
	(1.38 - 1.32)	(1.92 - 1.85)	(1.97 - 1.90)	(2.71 - 2.60)	(2.63 - 2.54)
R <sub>merge</sub>	0.124 (1.170)	0.167 (0.831)	0.143 (0.754)	0.217 (0.699)	0.137 (0.311)
<i>R</i> <sub>pim</sub>	0.043 (0.481)	0.078 (0.328)	0.114 (0.640)	0.088 (0.334)	0.065 (0.185)
Ι/ σΙ	40.0 (2.0)	13.0 (1.6)	6.9 (1.6)	8.3 (1.6)	11.6 (3.1)
Completeness (%)	95.0 (90.7)	99.7 (100)	97.5 (95.7)	99.4 (95.6)	98.8 (92.1)
Redundancy <sup>a</sup>	3.3 (3.1)	7.5 (7.4)	3.2 (3.2)	6.9 (4.8)	4.9 (3.1)
CC <sub>1/2</sub>	0.706 (0.215)	0.996 (0.895)	0.982 (0.480)	0.984 (0.703)	0.974 (0.882)
Refinement					
Resolution (Å)	36.26 - 1.32	17.06 - 1.86	50.55 - 1.90	49.71 - 2.60	47.22 - 2.54
	(1.37 - 1.32)	(1.92 - 1.86)	(1.97 - 1.90)	(2.69 - 2.60)	(2.63 - 2.54)
No. reflections	170694 (15998)	(59011/5804)	58345 (5700)	26488 (2368)	13596 (1131)
R <sub>work</sub> / R <sub>free</sub>	0.125/0.147	0.183/0.224	0.181/0.288	0.189/0.265	0.147/0.157
	(0.241/0.273)	(0.267/0.314)	(0.221/0.356)	(0.234/0.325)	(0.194/0.214)
No. atoms					
Protein	5733	5570	5572	5595	2778
Ligand/ion	83	52	52	80	27
Water	644	343	442	205	143
<i>B</i> factors (Å <sup>2</sup> )					
Protein	18	22	18	26	17
Ligand/ion	28	27	36	38	26
Water	30	28	27	26	20
R.m.s. deviations					
bond lengths (Å)	0.010	0.006	0.009	0.005	0.009
bond angles ( <sup>o</sup> )	1.2	0.9	1.1	0.9	1.1

#### Supplementary Table 4. Data collection and refinement statistics.

Each dataset was collected from a single crystal. \*Values in parentheses are for highest-resolution shell.

	HOZn <sup>2+</sup> distance (Å)	C=OZn <sup>2+</sup> distance (Å)
hCD2-TSA	2.0	2.3
	2.2	2.6
zCD1-TSA	1.9	2.9
	2.0	2.5
zCD2-TSA	2.2	2.4
zCD2-SAHA	2.1	2.2
	2.0	2.2
zCD2-Panobinostat	2.0	2.6
	2.1	3.5
zCD2-Belinostat	2.2	2.4
	2.1	2.6
zCD2-Oxamflatin	2.2	2.3

Supplementary Table 5. Hydroxamate inhibitor-Zn<sup>2+</sup> coordination interactions.



**Supplementary Figure 1. HDAC6 substrates 1-13 and inhibitors. (a)** Peptide substrates used for HDAC6 activity assays (AMC = aminomethylcoumarin). These commercially available substrates were selected based on their peptide sequence variations, which allowed us to probe HDAC6 substrate specificity. 1, Ac-Ser-Asp-Lys(Ac)-AMC, derived from α-tubulin. 2, Ac-Lys(Ac)-AMC, generic acetyllysine substrate. 3, Ac-Arg-His-Lys(Ac)-Lys(Ac)-AMC, derived from p53. 4, Arg-His-Lys-Lys(Ac)-AMC, derived from p53. 5, Ac-Gln-Pro-Lys-Lys(Ac)-AMC, derived from p53. 6, Ac-Lys-Gly-Gly-Ala-Lys(Ac)-AMC, derived from histone H4. 7, Ac-Gly-Ala-Lys(Ac)-AMC, derived from histone H4. 8, Ac-Arg-Gly-Lys(Ac)-AMC, derived from histone H4. 9, Ac-Gly-Ala-Lys(Ac), derived from histone H4. 10, Ac-Ala-Lys(Ac), derived from histone H4. 11, Ac-Ala-Lys(Ac)-Pro-NH<sub>2</sub>, artificial acetyllysine substrate. 12, Ac-Ala-Lys(Ac)-Ala-NH<sub>2</sub>, artificial acetyllysine substrate. 12, Ac-Ala-Lys(Ac)-Ala-NH<sub>2</sub>, artificial acetyllysine substrate. 13, Ac-Thr-Lys(Ac)-Pro-Ile-Trp-NH<sub>2</sub>, derived from Hsp90. (b) Inhibitors used in this study. Abbreviations: SAHA, suberanilohydroxamic acid; fl-SAHA, fluorescein-SAHA; HPOB, *N*-hydroxy-4-(2-[(2-hydroxyethyl)(phenyl)amino]-2-oxoethyl)benzamide.

8



Supplementary Figure 2. Enzyme activity measurements. Specific activities of human and zebrafish HDAC6 constructs assayed with fluorogenic substrates 1-8 and nonfluorogenic substrates 9-13 (substrates are illustrated in **Supplementary Fig. 1a**; data represent mean values  $\pm$  s.d. (n=3)). No activity is observed with hCD1 and Y386F/Y782F hCD12, so these constructs do not appear in the figure.



Supplementary Figure 3. Michaelis-Menten plots. Steady-state kinetics for human and zebrafish HDAC6 constructs assayed with fluorogenic substrates 1-8 and nonfluorogenic substrates 9-13 (substrates are illustrated in Supplementary Fig. 1a; data represent mean values  $\pm$  s.d. (n=3)).





Binding affinity of fl-SAHA determined by fluorescence anisotropy.<sup>a</sup>

Enzymes	<i>K</i> <sub>d</sub> (μΜ)
hCD12	$K_{\rm d1} = 0.2 \pm 0.1, \ K_{\rm d2} > 40^{\rm b}$
hCD12 H255V	0.24 ± 0.09
hCD12 H651V	30 ± 20
zCD1	1.6 ± 0.3
zCD1 H82F/F202Y	$2.9 \pm 0.7$
zCD2	0.05 ± 0.02
zCD2 S531A	0.17 ± 0.04

<sup>a</sup> Similar affinity trends are observed for inhibitor binding to the catalytic domains of human and zebrafish HDAC6, in that the fluorescent inhibitor binds to CD1 domains with lower affinity than to CD2 domains. While inhibitor affinity is lower for zCD1 compared with zCD2, it is not as low as that for hCD1 compared with hCD2. <sup>b</sup>  $K_{d2}$  cannot be accurately determined due to nonsaturating data.

**Supplementary Figure 4. Ligand binding to catalytic domains of hHDAC6. (a)** We probed inhibitor binding to hHDAC6 using fluorescence anisotropy spectroscopy and fl-SAHA (**Supplementary Fig. 1b**). Ligand binding is characterized by a biphasic curve indicating a high affinity site ( $K_{d1} = 0.2 \pm 0.1 \mu$ M) and a low affinity site ( $K_d > 40 \mu$ M) (data represent mean (n=3) ± s.d.). We then disabled ligand binding in each domain by mutating Zn<sup>2+</sup> ligand H255V in hCD1 and H651V in hCD2. H255V hCD12 retains high affinity ligand binding activity, whereas H651V hCD12 exhibits 150-fold diminished affinity. These results are consistent with high-affinity binding only to hCD2. (b) Binding affinities of fl-SAHA to human and zebrafish HDAC6 constructs (data represent mean ± s.e.m. (n=3)).



Supplementary Figure 5. Crystal structure of MBP-hCD2 complexed with maltose and TSA. (a) Simulated annealing omit map (grey mesh,  $3.0\sigma$ ) shows ligands bound to each protein. The linker (red) consists of the C-terminus of MBP (containing mutations D358A, E359A, K362A, D363A, R367N, and  $\Delta$ I368-K370) linked to an Ala<sub>3</sub> segment, which then connects to the N-terminus of hCD2. This engineered linker forms an  $\alpha$ -helix (red) that provides a complementary surface that facilitates MBP-hCD2 interactions. (b) Superposition of the hCD2-TSA (wheat; orange TSA) and zCD2-TSA (light blue; blue TSA) complexes. Residues D770-H771 in zCD2 are disordered and indicated by a dotted line on the right-hand side of the image. The Zn<sup>2+</sup> ion and K<sup>+</sup> ions in hCD2 are white and purple spheres, respectively; the corresponding metal ions in zCD2 are blue spheres. Only two active site residues differ between zCD2 and hCD2: N530 and N645 of zCD2 appear as D567 and M682 in hCD2, and these residues are located at the mouth of the active site, distant from the ligand binding site. (c) Molecular surface comparison of hCD2 (left) and zCD2 (right), color-coded by electrostatic potential (red to blue, -10 to 10 kT/e). TSA ligands are shown in ball-and-stick fashion to indicate the active site.



Supplementary Figure 6. Structure of zCD2. (a) Superposition of the zCD2-substrate 1 complex (light blue, dashed line indicating the disordered loop D770-H771) with unliganded zCD2 (wheat, dashed line indicating the disordered loop residues H771-L772). The substrate is indicated by a stick-figure; its N-terminal acetyl group and the serine side chain are not modeled due to disorder. The catalytic Zn<sup>2+</sup> ion is a white sphere and two K<sup>+</sup> ions are salmon spheres. Only minimal structural changes accompany substrate binding. (b) Superposition of the zCD2substrate 8 complex (light blue) with unliganded zCD2 (wheat, dashed line indicating the disordered loop residues H771-L772). The substrate is indicated by a stick figure; its N-terminal acetyl group and the arginine side chain are not modeled due to disorder. The catalytic  $Zn^{2+}$  ion is a white sphere and two K<sup>+</sup> ions are salmon spheres. Only minimal structural changes accompany substrate binding. (c) Simulated annealing omit map (blue mesh,  $3.0\sigma$ ) of the trifluoroketone inhibitor 7-[(3-aminopropyl)amino]-1,1,1-trifluoroheptan-2-one bound as a tetrahedral gem-diolate, which mimics the tetrahedral intermediate and its flanking transition states in catalysis (orange stick figure; F atoms are green); inhibitor O1 is closer to Zn<sup>2+</sup> than O2, likely due to electrostatic stabilization of O1 as the oxyanion by Zn<sup>2+</sup> coordination as well as the adjacent CF<sub>3</sub> group.

Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	1 MDAVPDTKGS SCHOOLS SCHOLL SGPPR GTPETK NNKKNL GEARRKGRMDRS KAEEEMSNE LONLDVQGKS-KATGTGL YVDAFTR 1 MTSTGQDSTTTRQRSRONPOSPDDS VTSKR SGPPR GGKGAVPHSIPNLAEVKKKGKKKLSG - PAEEDL VGLOGUDNE TRVPVGTGL V DEGLND 1 MTSTGQDSST-RQRKSRNNPOSPLODS ATLKR GGKKGAVPHSSPNLAEVKKKGKKKLSG - PAEEDL VGLOGUDNE TRVPVGTGL V FDGLND 1 MTSTGDDSTTPKERSSRNNPOSPLODS STTK SGKGAVPHSSPNLAEVKKKGKKKLSG - PAEEDL VGLOGUDNE TRVPVGTGL V FDGLND 1 MTSTGDDSTTPKERSSRNNPSPHDS STTK SGKGAVPHSSPNLAEVKKKGKKKLSG - PAEEDL VGLOGUDNE TRVPVGTGL V FDGLND 1 MTSTGDDSTTPKERSSRNNPSPHDS STTK SGKGAVPHSSPNLAEVKKKGKKKLSG - PAEEDL VGLOGUDNE TRVPVGTGL V FDGLND 1 MTSTGDDSTTPKERSSRNNPHSPHDS STTTS SGKGAVPHSSPNLAEVKKGKKKLSG - PAEEDL VGLOGUDNE STRVPYGTGL V FDGLND 1 MTSTGDDSTTPKERSSRNPHSPHDS STTTS SGKGAVPHSSPNLAEVKKGKKKLSG - PAEEDL VGLOGUDNE STRVPYGTGL V FDGLND 1 MTSTGDDSTTPKERSSRNPYSTHDS STTTS SGGLVLDDGLNE SGGLVGKGGGKGS SGKGGSGGSGGSGGGGGGGGGVG STRVFKFGGHKKLSG - NAEGDL VGLOGUDNE SGTVGTGAVF 1 MSSTGDDSTTVKERSKLOTTAGATAGGTGSSGS SGS SGS SG SGS SG SGSGG SGSGGGGGGG	FHC 76 FHC 99 FHC 98 FHC 98 FHC 99 FHC 99 FHC 99 TCC 122 HCC 119 HKC 117
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	77 WASHECARVSTVMEMLETEGLLGRCVSTCARAVTEDELLUHT KEVYELMKSTGONNET - EELKTLARKYD YVLHPGFFSSACLSVGSVLOLVOKWTSOLRIGFSIN PPOHLAOADKMN 100 LWDDSPEGPERLHAIREOLI LEGLLGRCVSTCARFAEKEELMLVHSLEYIDLMETTGYMNE - GELRVLARTYD VVLHPNSYSCACLASGSVLRLVDAVLGAEIRNGMAIIRPPOHLAOADKMN 100 LWDDSPESPERLHAIREOLI LEGLLGRCVSTCARFAEKEELMLVHSLEYIDLMETTGYMNE - GELRVLAGTYD VVLHPNSYSCACLATGSVLRLVDAVLGAEIRNGMAVI RPPOHLAOADKMN 100 LWDDSPESPERLHAIREOLI LEGLLGRCVSTCARFAEKEELMLVHSLEYIDLMETTGYMNE - GELRVLAGTYD VVLHPNSYSCACLATGSVLRLVDAVMGAEIRNGMAVI RPPOHLAOADKMN 100 LWDDSPESPERLHAIREOLI LEGLLGRCVSTCARFAEKEELMLVHSLEYIDLMETTGYMNE - GELRVLAGTYD VVLHPNSYSCACLATGSVLRLVDAVMGAEIRNGMAVI RPPOHLAORSKMN 100 LWDDSPESPERLHAIREOLI LEGLLGRCVSTCARFAEKEELMLVHSLEYIDLMETTGYMNE - GELRVLAGTYD VVLHPNSYSCACLATGSVLRLVDAVMGAEIRNGMAVI RPPOHLAORSKMN 100 LWDDSPESPERLHAIREOLI LEGLLGRCVSTCARFAEKEELMLVHSLEYIDLMETTGYMNE - GELRVLAGTYD VVLHPNSYSCACLATGSVLRLVDAVMGAEIRNGMAVI RPPOHLAORSKMN 100 LWDDSPESPERLHAIREOLI LBCULDRCVSTCARFAEKEELMLVHSLEYIDLMETTGYMNE - GELRVLAGTYD VVLHPNSYTCACLASGSVLRLVDAVLGAEIRNGMAII RPPOHLAORSKMN 100 LWDDSPESSERLH AIREOLI LBCULDRCVSTCARFAEKEELMLVHSLEYIDLMETTGYMNE - GELRVLAGTYD VVLHPNSYTCACLASGSVLRLVDAVLGAEIRNGMAII RPPOHLAONSKM 100 LWDDSPESSERLH IKEOLI DRCVSTCARFAEKEELMLVHSLEYIDLMETTGYMNE - GELRVLADTYD VVLHPNSYTCACLASGSVLRLVDAVLGAEIRNGMAII RPPOHLAONSKM 100 LWDDSPESSERLH IKEOLI DRCVSTCARFAEKEELMLVHSLEYIDLMETTGYMNE - GELRVLADTYD VVLHPNSYTCACLASGSVLRLVDAVLGAEIRNGMAIIRPPOHLAONSKM 100 LWDDSPESSERLH IKEOLI DRCVSTGARFAEKEELMLVHSLEYIDLMETGYMNE - GELRVLADTYD VVLHPNSYTCACLASGSVLRLVDAVLGAEIRNGMAIIRPPOHLAONSKM 100 LWDSPESSERLH IKEOLI DRCVSTGARFAEKEELMLVHSLEYIDLMETGYME - GELRVLADTYD VVLHPNSYTCACLASGSVLRLVDAVLGAEIRNGMAIIRPPOHLAONSKM 100 LWDSFPESSERLH IKEOLI DRCVSTGARFAEKEELMLVHSLEYIDLMESTGUNNUVELELASSTDICALASGSVILL VVLWELREIRNGLAVHAN IN HANNEN CH	SFC 203 SYC 226 GYC 225 GYC 225 GYC 226 GYC 226 GYC 226 GYC 249 GYC 247 GYC 245
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	204 MFINLA IAAR YAQKHRVQRVL VWW VHGGEIQYI FEDPSLYFS WRYEDS SWRYEDS SVBSGAGGYN IN LPWNKYGMESCDYI TAFQOLLLPAAFFALGO 227 MFINLA VAAR YAQKHRIGRYL VWW VHGGEIGYI FOODPSLYFS IR YEHRFWHLKEDSS VSSGAGGYN INVWWGGRDADYI AAFLHULPYAEFGOLLLPAAFFALGO 228 MFINLA VAAR YAQKHRIGRIL VWW VHGGEIGYI FOODPSLYFS IR YEHRFWHLKSNWSTIGFGGGGYT INVPWNGGRDADYI AAFLHULPYAEFGOLLLAAGFFALGO 228 MFINLA VAAR YAQKHRIGRIL VWW VHGGEIGYA FOODPSLYFS IR YEHRFWHLKSNWSTIGFGGGGYT INVPWNGGRDADYI AAFLHULPYAFFGOLLLAAGFFALGO 229 MFINLA VAAR YAQKHRIGRIL VWW VHGGEIGYA FOODPSLYFS IR YEHRFWHLKSNWSTIGFGGGGYT INVPWNGGRDADYI AAFLHULPYAFFGOLLLAAGFFALGO 220 MFINLA LAAR YAQCHOIGYL VWW VHGGEIGYA FOODPSLYFS IR YEGRFWHLKSNWSTIGFGGGGYT INVPWNGGRDADYI AAFLHULPYAFFGOLLLAAGFFALGO 227 MFINYA VAAR YAQKHOIGYL VWW VHGGEIGYA FOODPSLYFS IR YEGRFWHLKSNWSTIGFGGGGYT INVPWNGGRDADYI AAFLHULPYAFFGOLLUAAGFFALGO 250 MFINLA LAAR YAQCHOIGYL VWW VHGGEIGYA FOODPSLYFS IR YEGRFWHLKSNWSTIGFGGGGYT INVPWNGGRDADYI AAFLHULPYAFFGOLULAAGFFALGO 250 MFINLA LAAR YAQCHVIGW VHGGTGFLF FNDPSLYFS IR YEGRFWHLKSNWSTIGFGGGGYT INVPWNGGRDADYI AAFLHULPYAFFGOLULAAGFFALGO 250 MFINLA LAAR YAQCHVIGW VHGGTGFLF FNDPSLYFS IR YEGRFWHLKSNWSTIGFGGGGYT INVPWNGGRDADYI AAFLHULPYAFFGOOLULAAGFFALGO 250 MFINLA LAAR YAQCHVIGW VHGGTGFLF INDPSLYFS IR YEGRFWHLKSNWSTIGFGGGGYT INVPWNGGRDADYI AAFLHULPYAFFGOOLULAAGFFALGO 250 MFINLA LAAR YAQCHVIGW VHGGTGFLF INDPSLYFS IR YEGRFWHLKSNWSTIGFGGGWYT INVPWNGGRDADYI AAFLHULPYAFFGOOLULAAGFFALGO 250 MFINLA LAAR YAQCHVIGW VHGGTGFLF INDPSLYFS IR YEGRFWHLKSNWSTIGFGGGWYT INVPNGLAGRADADYI HVFLNILLFI IL YAFFGOOLULAAGFFALGO 250 MFINLA LAAR YAQCHVIGW VHGGTGFLF INDPSLYFWYNG VYSI IR YEGRFWHLKSNWSTGFGGGWYT INVPNGLWGRDADYI IAAFLHULPYAFFGOOLULAAGFFALGO 250 MFINLA LAAR YQCHVIGGTIFF VUDDY VYSI IR YEGRFWHLKSNSTGFGGGWYT INVPNGLAGRADADYI HVFLNILLFI IL YAFFGOOLULAAGFFALGO 260 FFNUKA TGGLUGW HGGTGFLF VUDDY VYSI IR YEGRFWHLKSNSTGFGGGWYT INVPNGLWGRDADYI INVFNGLWGRDADYI IN LY IAAFGALG 246 FFNUKA TGGLUGW HGGTGFLF VUDDY VYSI IR YEGRFWHLKSNSTGFGGGWYT INVPNGLWGRDADYI INVFNGLWGRDADYI IN LIGGUN HYGGUN HGGGTGFLFI INVFNGLWGWGDADYI IN	P - 329 P - 352 P - 351 P K V 353 P - 352 P - 352 P - 352 P - 375 P - 373 P - 371
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	330 K GOMOVSBECJSIL THILIKOVA ORALVLA EGGUNLOSTAEGUCASMASLLODPOP. HLBSGAPOEBALKSISKTISDUYPFWKSLOTFEGG PLSEVSPLP. 351 K GEMAATPAGFAU THILIKOLAGKILISLEGGUNLAALKGUSASLHTILLODPOP. HLBSGAPOEBALKSISKTISDUYPFWKSLOTFEGG PLSEVSPLP. 352 K GEMAATPAGFAU THILIKOLAGKILISLEGGUNLAALKGUSASLHTILLODPOP. HLBSVOYOASAGTSIYCTLEALEPFWEVLRESVETGE. EDEVEEAVLEEEE 353 K GEMAATPAGFAU THILIKOLAGKILISLEGGUNLAALKGUSSLHTILLODPOP. HLBSVOYOASAGTSIYCTLEALEPFWEVLRESVETOE. EDEVEEAVLEEEE 354 ROEMAATPAGFAU THILIKOLAGKILISLEGGUNLABLAEGUSSLHTILLODPOP. HLBSVOYOASAGTSISCTLEALEPFWEVLRESVETOE. EDEVEEAVLEEEE 355 K GEMAATPAGFAU THILIKOLAGKILISLEGGUNLABLAEGUSSLHTILLODPOP. LESPOAPOPSAASLSCTLEALEPFWEVLRESVETOE. EDEVEEAVLEEEE 356 K GEMAATPAGFAU THILIKOLAGKILISLEGGUNLESLEGGUNLABLAEGUSSLHTILLODPOP. LESPOAPOPSAASLSCTLEALEPFWEVLRSVETELEDEDTVEGDUEEEE 376 K GEMAATPAGFAU THILIKOLAGKILISLEGGUNLESLEGGUNLESLEGUNLESUSASLHTILLODPOP. LEFPOAPOPSAASLSCTLEALEPFWEVLVSSELEKEDTVEKDOVEEKE 376 K GEMAATPAGFAU THILIKOLAGKILISLEGGUNLESLEGGUNLESLEGUCACLALLEDPOP. LEFPOAPOPSAASLSCTLEALEPFWEVLVSSELEENDTVEKDOVEEKE 377 K EDEWEVTPÄCYPHLINDLIKLADARVAVVEGOUNDESLEGUCACLAALLODPOP. LEFPOAPOPSAASLSCILDSISDTUSAHCRIKKVLODVESEGEAIDEPOSAEEOPPOA 374 EDEWEVTPÄCYPHLINDLIKLADARVAVVEGOUNDESLEGUCACLAALLEDPOPP. VETVPLIPSAFOOSALDISISDTUSAHCRIKKVLODVESEGEANDEFOEDHVEHVIKSAET 375 K EDEMAATPAGFAULIPHLIKLADARVAVVEGOUNDESLEGUCACLAALLEDPOPP. VETVPLIPSAFOOSAALDISISDTUSAHCRIKKVLODVESLEKEND 376 EDEMEVTPÄCYPHLINDLIKLADARVAVVEGOUNDESLEGUCACLAALLEDPOPPLVETVPLIPSAFOOSAALDISISDTUSAHCRIKKVLODVESLEKENDTAVENUUNGAPPPORA 377 EDEMENTATORAULIPHLIKLADARVAVVEGOUNDESLEGUNGEN LEGUCACLAALLEPPOPPLVETVETVITACYPHUKSLONAVEDUNTASLEDEDTVEKDHVENDA 378 EDEMENTATORAUVENTEGOUNDESLEGUCKSSENTILENDITUSTIK	430 459 458 461 460 485 485 - YP 498 VYE 496
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	431	SN 536 CP 573 CP 572 CP 575 CS 574 CP 572 NS 602 LHP 626 LHP 624
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	337 ESYTCALLAASSCENSAQAILTEQVENAVA IVEPPOTIAEKÜTAGEGEFNTAALTARVAGSITESLEVLIVOOVYHENGTOH FEEDSULYISLHEVEDAFFPNSEDANVDKVGLOK 574 STFACALATGAACRLVEAVISEUVIGAAVVEPOTIAEQDAGEGEFNSVAVARHAQTISGAILEVLIVOOVYHENGTOH FEEDSULYISLHEVEDAFFPNSEDANVDKVGLOK 575 STFACALATGAACRLVEAVISEUVIGAIVVEPOTIAEQDAGEGEFNSVAVARHAQTISGALEVLIVOOVYHENGTOH FEDDSULYUSLHEVDHOTFFPMGDEGASSOVGRAA 576 STFACALATGAACRLVEAVISEUVIGAIVVEPOTIAEQDAGEGEFNSVAVARHAQTISGALEVLIVOOVYHENGTOH FEDDSULYUSLHEVDHOTFFPMGDEGASSOVGRAA 575 STFACALATGAACRLVEAVISEUVIGIAIVEPOTIAEQDAGEGEFNSVAVARHAQTISGALEVLIVOOVYHENGTOH FEDDSULYUSLHEVDHOTFFPMGDEGASSOVGRAA 575 STFACALATGAACRLVEAVISEUVIGIAIVEPOTIAEQDAGEGEFNSVAVARHAQTISGALEVLIVOOVYHENGTOH FEDDSULYUSLHEVDHOTFFPMGDEGASSOVGRAA 575 STFACALAGAACRLVEAVIAEUVIGIAIVEPOTIAEQDAGEGEFNSVAVARHAQAISGHALEVLIVOOVYHENGTOH FEDDSULYUSLHEVDHOTFFPMGDEGASSOVGRAA 575 STFACALAGAACRLVEAVIAEUVIGIAIVEPOTIAEQDAGEGEFNSVAVARHAQAISGHALEVLIVOOVYHENGTOH FEDDSULYUSLHEVDHOTFFPMGDEGASSOVGRAA 576 STFACALAGAACRLVEAVIAEUVIGIAVVEPOTIAEQDAGEGEFNSVAVARHAQAISGHALEVLIVOOVYHENGTOH FEDDSULYUSLHEVDHOTFFPMGDEGASSOVGRAA 576 STFACALAGAACRLVEAVIAEUVIGIAVVEPOTIAEQDAGEGEFNSVAVARHAQAISGHALEVLIVOOVYHENGTOH FEDDSULYUSLHEVDHOTFFPMGDEGASSOVGRAA 576 STFACALAGAACRLVEAVIAEUVIGIAVVEPOTIAEQDAGEGEFNSVAVARHAQAISGHALEVLIVOOVYHENGTOH FEDDSULYUSLHEVDHOTFFPMGDEGASSOVGRAA 577 STFACALAGAACRLVEAVIAEUVIGIAVVEPOTIAEQDAGEGEFNSVAVARHAQAISGHALEVLIVOOVYHENGTOH FEDDSULYUSLHEVDHOTFFPMGDEGASSOVGRAA 578 STFACALAGAACRLVEAVIAEUVIGIAVVEPOTIAEQDAGEGEFNTAAIAAGAISGHA	RG 661 TG 698 G IG 697 G TG 700 G TG 699 G TG 699 G TG 697 G AG 730 G RG 750 G E G 748
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	662 YUN I PWNGK KIGDPE YMARFH LVWP ARF FABELV LVSAGFOAR COPLOG COVSPECANEL THOLKSLAG RVLIILEGOVNITES SEMSMOTSMULGDSPESIDHET PLKTS   689 FUN VANORPRIGOADVLAAMHELV I AV FENELV LVSAGFOAR COPLOG COVSPECANEL THOLKSLAG RVLIILEGOVNITES SEMSMOTSMULGDSPESIDHET PLKTS   698 FUN VYMORPRIGOADVLAAMHELV I AV FENELV LVSAGFOAR COPLOG COVSPECANEL THULKGLAG RILLILEGOVINTES SEMSMOTSMULGDSPESIDHET PPLSG   698 FUN VYMORPRIGOADVLAAMHELV I AV FENELV LVSAGFOAR COPLOG COVSPECANEL THULKGLAG RILLILEGOVINTES SEMACTHSLUD PPPLITER PPLSG   698 FUN VYMORPRIGOADVLAAMHELV I AV FENELV LVSAGFOAR COPLOG COVSPECANEL THULKGLAG RILLILEGOVINTES SEMACTHSLUD PPPLITER PPDSG   701 FUN VYMORPRIGOADVLAAMHELV I PLAY FENELV LVSAGFOAR COPLOG COVSPECANEL THULKGLAG RILLILEGOVINTES SEMACTHSLUD PPPLITER PPOSG   700 FUN VANOPRIGOADVLAAMHELVIP AV FENELV LVSAGFOAR COPLOG COVSPECANEL THULKGLAG RILLILEGOVINTES SEMACTHSLUD PPPLITER PPOSG   700 FUN VANOPRIGOADVLAAMHELVIP AV FENELV LVSAGFOAR COPLOG COVSPECANEL THULKGLAG RILLILEGOVINTES SEMACTHSLUD PPPLITER PPLSR   731 NVN I PWNGKKIGO PE LVAAHHELVIP AV FENELV LVSAGFOAR COPLOG COVSPECANEL THULKGLAG RILLILEGOVINTES SEMMOTHSLUD PPPLITER PPLSR   731 NVN I PWNGKKIGO PE LVAAHHELVIP AV FENELV LVSAGFOAR COPLOG COVSPECANETHLIKGLAG RILLILEGOVINTES SEMMOTHSLUD PPT TSOLR PPLSR   731 NVN I PWNGKKIGO PE LVAAHHELVIP AV FENELVELVSAGFOAR COPLOG COVSPECANETHLIKGLAG RILLILEGOVINTES SEMMOTH	A T V 781 A L A 818 A L V 817 A L A 820 A Q A 819 A L A 817 A F N 850 C V E 878 A V H 868
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	782 II NIVURAHAPPINS LEVNIPESIRLELPSPKPKKKCTPGGKGKKSPROSTPSOSGQATHRPDSOGGATHRPDSOGGATLACI 818 II TEI DVIRRYWRELRUMKEDREGPSSKLVIKKAPO KRRLAENTTREKKVLEAGMKVTSASFGEETPOGINSETAVVAITGDOPSEAATGGATLACI SEAAIGGAMLGGTTSEE 818 ISEVI DVIRKYWRELRUMKEDKEERSSSKIVIKKIPPIASPVSAKEMITPKGKVLEAGMKKPTALLPKESTLOGAKSKMAAAVLAGOOSEGAAKG. 821 ISEVI DVIRKYWRELRUMKEDKEERSSSKIVIKKIPPIASPVSAKEMITPKGKVLEAGMKKPTALLPKESTLOGAKSKMAAAVLAGOOSEGAAKG. 820 IT TLIDVIRRYWRELRUMKEDKEERSSSKIITKPSOPASPGLANVITTYKGNVLEAGMKKPTALLPKESTLOGAKAKALAVLAGOOSEGAAKG. 830 IT TLIDVIRRYWRELRUMKKNKEERSSSKIITKPSOPASPGLANVITTYKGNVLEAGMKKPTALLPKESTLOGAKAKALAVLAGOOSEGAAKG. 840 IT TLIDVIRRYWRELRUMKKNKEERSSSKIITKPSOPASPGLANVITTYKGNVLEAGMKPTALPKESTLOGAKAKALVELTGOOSDTATO. 851 VKKVROAMRKYMRSLRUNAPGGAHNLOHSPVNNSOEGRFATEAGPSAMISPOGASPPARIFGATSVSVKVSPPOGITSEALVELTDOOSSKAAT 851 VKKVROAMRKYMRSLRUNAPGGAHNLOHSPVNNSOEGRFATEAGPSAMISPOGASPPARIFGETISVSVKVSVFVSKTEPPOTUSEALVELTDOOSSKAAT 851 VKKVROAMRKYMRSLELNVAPGGAHNLOHSPVNNSOEGRFATEAGPSAMISPOGASPPARIFGETIFLOGAKSKRVKVKVKTDOVLAENTSVSVKVVKVKTAGVTARAVELTDOOSSKAAT 851 VKKVROAMRKYMRSLELNVAPGGAHNLOHSPVNNSOEGRFATEAGPSAMISPOGASPPARIFGETIFLOGORSGKSKFVKVKVKTDOVLAELSVSVKVSPTORTSKAALVELTDOOSSKAAT 851 VKKVROAMRKYMRSLELNVAPGGAHNLOHSPVNNSOEGRFATEAGPSAMISPOGASPARFFVKVKVKVKVKVKVKVKVKVKVKVKVKVKVKVKVKVKV	AVG 944 916 919 917 914 951 973 968
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	854 -ADLLSLNLINTSTSSNTSPESVÄG-GARRKVKPNTHRVSAGLESESF-LKLN -PEKAGADDTDVK-AETPNRSISEVVTODSSAVK   945 GATPDQTTSEETVGGAILDDTSEKAVGGATLADTSEAAMEGATLDOTTSE PAPOGTELIOTELASSTDHATPPTSPVOGTTPOISPSTLISELTLESESAGSEG   947 GATPDQTTSEETVGGAILDDTSEKAVGGATLADTSEAAMEGATLDOTTSE PAPOGTELIOTELASSTDHATPPTSPVOGTTPOISPSTLISELTELSESAGEASES   947 TTLDLATSKETVGGATTDLWASAAPENPRNOTTS.VEALGETEPTTPASHTNKOTTGASPLOGVTAQSIGLQUSTLELSEAEEAMDES   940 TTLDLATSKETVGGATTDLWASAAPENPRNOTTS.VEALGETEPTTPASHTNKOTTGASPLOGVTAQSIGLQUSTLELSEAEEAMDES   940 TTLDLATSKETVGGATTDLWASAAPENPRNOTTS.VEALGETEPTTPASHTNKOTTGASPLOGVTAQSIGLQUSTLELSEAEEAMDES   940 TTLDLATSKETVGGATTDLWASAAPENPRNOTTS.VEALGETEPTTPASHTNKOTTGAAPLOGVTAGSIGLGUSTLELSEAEEAMDES   941 GAALDQTISECATGAAELIONPASCINNEIPPALPVOGATAGISPSKLMANIERILDLDSTTGEPSE   945 GAALDQTISECATGAAELIONPASCINNEIPPALPVOGATAGISPSKLMANIERILDLDSTTGEPSE   945 GAALDQTISECATGAAELIONPAGCINNEIPPALPVOGATAGISPSKLMANIERILDLDSTTGEPSE   945 GAALDQTISSAVGEPEPEPAITESOIMENTERVSCOSTGTGELMGAEERAGATGEKOVEP   952 PFTDPLAAPPHVMPTEEVSPAKLTSPKGTSGALSVAGEPEPEPAITESOIMENGAGGONGTPKSPTSTGTGTELMGAEGRADTAPERGATGEKKOVT   954 GAAVYPLKTOPHLRLIRPEAPRSLSGAESVGSTGEWVCLSC   956 GAAVYPLKTOPHLRLIRPEAPRSLSGAESVGSTGEVVCSGSCGONG   969 GAAUNTENTSAVAGAAUNCLSC	D 938 APG 1072 1006 1012 984 972 T - 1048 1022 1021
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	939 DESIVUE RAAGGAUTVOSVLERVFGAAATSVDTWV VODE PROPHLES VEPV RAGE DVY OPPEGE RGGE AENNI CUE PXVUE GRV VIGAV VIE RABLEVVEVA SE SYVHAVU 1073 EEN LGEAAGGADMAD.SMLMGGSRGLTDGAIFYAVTE PVT I SMOPHLMAVCPI FAAGLDV TOPGOV GSTVGVVG SVVVA MVCHEABEHFUL SVVD STVCVA SVVHAM 1071 EEGLLGEAAGGADMAS.LMLTGGFSDFNT GDVFTAVTPLSMOPHLMAVCPI FAAGLDV SOPGODGTIG SVMVGLTVGVVS SVVNA MVCHEABEHFUL SVD STVCV COAVVHED 1012 EEGLLGEAAGGADMAS.LMLTGGFSDFNT GDVFTAVTPLSMOPHLMAVCPI FAAGLDV SOPGODGTIG SVMVGLTVGVVS SVVNA MVCHEABEHFUL SVD STVCV COAVVHED 1013 EEGLLGEAAGGADMAS.LMLTGGFSDFNT GDVFTAVTPLSMOPHLMAVCPI FAAGLDV SOPGOTGTUG SVMVGLTVGVVS SVVNA MVCHEABEHFUL SVD STVCV COAVVHED 1014 EEGLLGEAAGGADMAS.LMLTGGFSDFNT GDVFTAVTPLSMOPHLMAVCPI FAAGLDV SOPGOTGTUG SVX00 SVVGLTVGVVS SVVNA MVCHEABEHFUL SVD STVCV COAVVHED 1015 EEGLLGEAAGGADMAS.LMLTGGFSDFNT GDVFTAV FPLSMOPHLMAVCPI FAAGLDV SOPGOTGTUG SVX00 SVCVTS SVVNA MVCHEABEHFUL SVD STVCV COAVVHED 1016 EEGLLGEAAGGADMAS.LMLTGGFSDFNT GDVFTAV FPLSMOPHLMAVCPI FAAGLDV SOPGOTGTUG SVX00 SVCVVCS SVCVVCGV IN AMVCHEABEHFUL SVD STVCV COAVVHED 1017 EEGLLGEAAGGADMAS.LMLTGGFSDFNT GDVFTAVFPLSVVS PLAVDPHLGAVGAV FAAGDAV SOPGOTGTUG SVX00 SVCVCS SVCVVCGV IN AMVLCHEGSGIPAUL SVAD STVCV COAVVHED 1018 EEGLLGEAAGGADMAS.LMLTGGFSDFNT GDVFTAVFPLGVSVSVVFAASIDV SSVCVCS SVCVCGVCGV IN AMVLCHEGSGIPAUL SVAD STVCV COAVVHED 1019 KELLEBAAGGSAGANNS.LMLTGGFSDFNT SVCVCVGSVCVCS SVCVCGVCGV IN AMVLCHEGSGIPAUL SVAD STVCV COAVVHED 1019 KELLEBAAGGSAGANNS.LMLTGGSSGFAVTEPPVDFIGSVSVVFASIDV SSVCVCS SVCVCGSVVCGVCGVVCGVCGVVVHCVS SVCVCSSVCVCGVCGVVVHENC 1024 KILLEBAAGSSAGANNS.LMLTGGSSGFAVTEPPVDFIGSVSVVVASIDV SSVCVCSSVCVCGSVVVHCUSSVCVCGSVVVHCUSSVCVCGSVVVHCNS SVCVCGVCGVVHHA 1025 KILLEBAAGSSAGANNS.LMLTGGSSGFAVTEPPVDFIGSVSVVASIDVCVSSVCVCSSVCVCSSVCVCGSVVHAN 1026 KILLEBAAGSSAGANVCHEGSSGFAVTEPVDFIGSVSVVVASIDVCVSSVCVCSSV	LHE 1063 LD1197 LQD1131 LQD1137 LLA1111 LLA1111 LLA1099 LFP1176 1120 1120
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	1064 AKNAAH LVKFGEG. IHP FN. 1198 VKN AHONKFGED. MPHPH. 1138 VKNAAHONKFGED. MPHSH. 1138 VKNAAHONKFGED. MPHSH. 1139 VKNAAHONKFGED. IPHSD. 1100 VKN I AHONKFGED. VPHSR. 1107 XKS AKYSSKEGANSGLECSGEGSAHGTPKSPPSTGSGRVLMGELEAAD I APOPEGDPLDVLSPGQR I GSGKEGKAPGPPSSTPGERSLEG I LKDLOLSDHAGNNPT I AFANETHLDPVGGARRKE	1081 1215 1149 1155 1129 1117 F I S 1304 
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly Bee	1305 TGRDAREEEQKEGRS I QAKGVTPTKQELLGEAAGGSEAGNPH I GFETLMESMANESGEGFAVTPLPWCPHLGSVSAVPPAGLDVRQLCAQCASELENWVCLTCYQVLCGRYVSQHMLCHGLASGH	   H L V 143; 
Zebrafish Human Mouse Rat Bovine Porcine Frog Fly	1433 LSFSDLSVWCYGCESYVHHQVQRGIVGGLERVPGSLLLIALGSLSLTPRPCSRPNLLPTAPSLGRKCNRCRGC	150

**Supplementary Figure 7. Sequence alignment of HDAC6 orthologues.** In each domain, the Zn<sup>2+</sup> ion binding residues are highlighted in red, the conserved catalytic residues in the active site are highlighted in orange, the aspartate residue interacting with the backbone NH groups flanking the acetyllysine substrate (D101 in HDAC8) is highlighted in sky blue, the newly identified conserved serine residue important for substrate binding in HDAC6 is highlighted in magenta, and the conserved buried active site cysteine residue is highlighted in green. Remaining conserved residues with are highlighted in blue. Sequence alignment prepared with Clustal Omega, figure prepared with Jalview.



#### Supplementary Figure 8. Macrocyclic HDAC inhibitors, and structure of the zCD2-HC

**toxin complex. (a)** A class of cyclic tetrapeptides containing 2-amino-8-oxo-9,10epoxydecanoic acid (L-Aoe, red), which partially mimics the HDAC substrate N<sub>E</sub>-acetyl-L-lysine. Both the terminal epoxide and adjacent carbonyl group of the Aoe side chain are essential for inhibitory potency against HDACs. **(b)** Superposition of the zCD2-HC toxin complex (light blue, white surface) with unliganded zCD2 (green) shows that the binding of HC toxin (orange stick figure) to zCD2 does not trigger any significant conformational changes. The Zn<sup>2+</sup> ion, K<sup>+</sup> ion, and water molecules are shown as white, salmon, and red spheres, respectively. Disordered segments are shown as dashed lines. Notably, this is the first crystal structure of HC toxin, and hence the first observation of the novel cis-trans-cis-trans peptide bond architecture of the cyclic tetrapeptide. **(c)** Close-up view of HC toxin binding, showing the solvent-accessible surface of the active site contour. Metal coordination and hydrogen bond interactions are shown in red and black dashed lines, respectively.



**Supplementary Figure 9. Crystal structure of zCD1. (a)** The overall structure of the zCD1-TSA complex (light blue) is highly similar to that of the zCD2-TSA complex (green), with an r.m.s. deviation of 0.58 Å for 311 C $\alpha$  atoms. The simulated annealing omit map (blue mesh, 3.0 $\sigma$ ) indicates TSA (orange) bound in the active site of zCD1; Zn<sup>2+</sup> ions are shown as white spheres. **(b)** Close-up view of the zCD1 active site. Metal coordination and hydrogen bond interactions are shown as red and blue dashed lines, respectively.



**Supplementary Figure 10. Structures of zCD2 complexed with pan-HDAC inhibitors and HDAC6-specific inhibitor HPOB.** Simulated annealing omit maps (blue mesh, 3.0σ) show each bound inhibitor. Metal coordination and hydrogen bond interactions are shown as red and black dashed lines, respectively; Zn<sup>2+</sup> ions are shown as white spheres. Inhibitors are shown as ocher stick figures and are as follows: (a) TSA; (b) SAHA; (c) Belinostat; (d) Panobinostat; (e) Oxamflatin; and (f) HPOB. The cap region of HPOB interacts with an adjacent monomer (wheat helix) in the crystal lattice.



Enzyme	MBP-	zCD1					zCD2			
	hCD2									
Inhibitor	TSA	TSA	TSA	HPOB	HC	Trifluoro-	Belinostat	Panobinostat	SAHA	Oxamflatin
					toxin	ketone				
IC <sub>50</sub> (nM)	3.9	7	1.4	1.8	600	1400	1.9	2.5	1.1	7.5
	± 0.8	± 1	± 0.2	± 0.3	± 90	± 200	± 0.3	± 0.5	± 0.3	± 0.9
Ki	2.2	4.8	0.8	1.0	350	800	1.1	1.4	0.6	4.3
(nM)	± 0.4	± 0.7	± 0.1	± 0.2	± 50	± 100	± 0.2	± 0.3	± 0.2	± 0.5

Supplementary Figure 11. Inhibition of HDAC6 by inhibitors used in crystal structure

**determinations.** Data were analyzed by logistic regression for IC<sub>50</sub> determination and the inhibition constant K<sub>i</sub> was calculated based on the Cheng-Prusoff equation assuming competitive inhibition,  $K_i = IC_{50}/(1+[S]/K_M)$ , as described in the Methods section. The enzyme concentrations used in this study were 3 nM for zCD2, 6 nM for MBP-hCD2, and 12 nM for zCD1. Since some IC<sub>50</sub> values are near the measurement threshold of [enzyme]/2, these values should be regarded as upper limits since the actual IC<sub>50</sub> may be even lower. Data represent mean  $\pm$  s.e.m. (n=3).



**Supplementary Figure 12**. Representative LC-MS traces (ESI-) of enzymatic reactions with substrate **9**. Due to its small size and polarity, the enzymatic product was dansylated in order to be efficiently separated from the reaction mixture for MS characterization. The enzymatic product can be quantified by comparing the integrated peak area with that of dansylamide ( $t_R$  = 2.40 min) as an internal standard to determine specific activity. The full mass spectrum of the dansyl-derivatized product peak (\*) is shown below.



**Supplementary Figure 13**. Representative LC-MS traces (ESI+) of enzymatic reactions with substrate **13** (corresponding to residue T2930-W297 of Hsp90). Full mass spectra of substrate (I) and product (II) peaks are shown below. The enzymatic product can be quantified by comparing the integrated peak area with that of an authentic synthetic product peptide Ac-Thr-Lys-Pro-Ile-Trp-NH<sub>2</sub> as an external standard to determine specific activity.