Loop Interactions and Dynamics Tune the Enzymatic Activity of the Human Histone Deacetylase 8

Micha B. A. Kunze[1], David W. Wright[2], Nicolas D. Werbeck[1], John Kirkpatrick[1], Peter V. Coveney[2,*] and D. Flemming Hansen[1,*]

[1] Institute of Structural and Molecular Biology, University College London, Gower Street, London WC1E 6BT. [2] Centre for Computational Science, Department of Chemistry, University College London, Gordon Street, London WC1H 0AJ, UK.

Supporting Information. Supporting material comprises two videos, Video S1-S2 and Figure S1-S5.

SUPPORTING VIDEOS

Video S1: Video S1 shows the L1-L2 loop region during the molecular dynamics simulation of the free HDAC8. The states as characterized in the main text are indicated during the video. The simulation structure is shown in cyan, which is overlaid with the 'binding state' structure of the crystal structure PDB 2V5W shown (green) and the crystal structure of PDB 1T64 (red). These overlays show how the L1 loop structures differ.

Video S2: Video S2 shows the L1-L2 loop region during the molecular dynamics simulation of the HDAC8:SAHA complex. SAHA is rendered as a cyan licorice representation. The simulation structure (cyan) is overlaid with the crystal structures of PDB 2V5W (green) and PDB 1T64 (red).

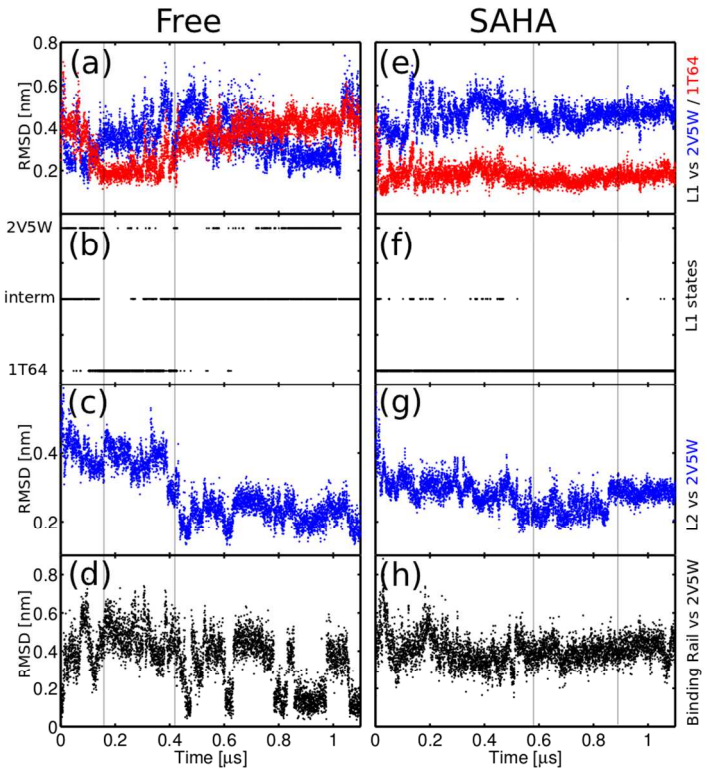

Figure S1. Comparison of the L1 and L2 loop conformation during the simulations with the corresponding conformation of the crystal structures. The left panel shows data for the free HDAC8 and the right panel data for the SAHA:HDAC8 complex. **(a)** and **(e)**: RMSD of the L1 loop against the binding state structure, PDB: 2V5W is shown in blue; RMSD of the L1 loop against the conformation observed in 1T64 shown in red, both RMSD calculated after a least squares fit of the rigid protein backbone. **(b)** and **(f)**: State of the L1 loop. Assignment as 2V5W (1T64) where the L1 loop conformation has an RMSD less than 2.6 Å with respect to the L1 conformation of 2V5W (1T64). An intermediate state was assigned if neither the 2V5W or the 1T64 state could be assigned. **(c)** and **(g)**: RMSD of the entire L2 loop against the binding state structure published by Vannini et al. in the structure PDB: 2V5W; after least squares fit of the rigid protein backbone. (**d)** and **(h)**: RMSD of the backbone of the binding rail residues Tyr100 and Asp101 after a least squares fit of the rigid protein backbone. When the RMSD value is below 0.25 nm in (**d**) the 'in' conformation of the binding rail, as found in 2V5W, is sampled.

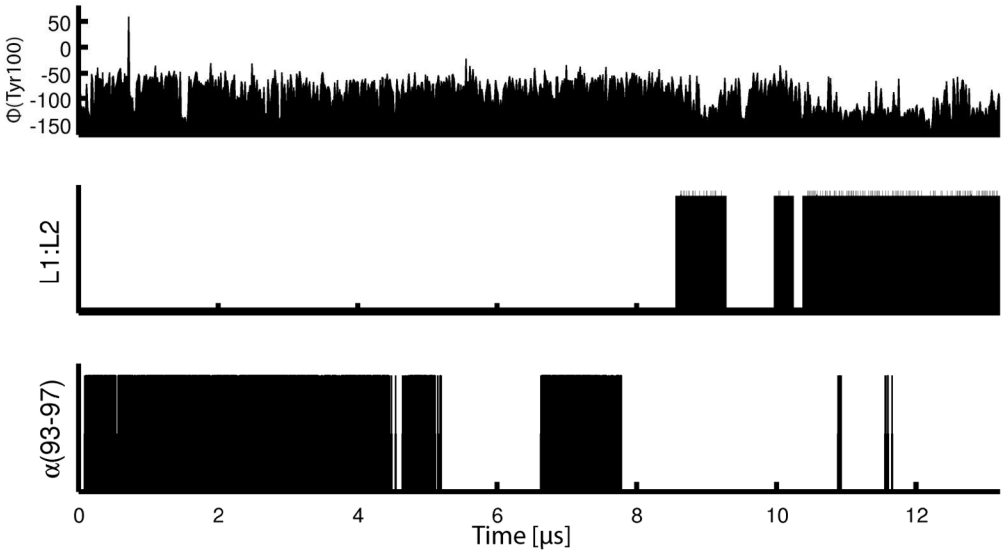

Figure S2. The three microkinetic states as a function of simulation time, as described in the main text, for the 13.2 μs long Anton simulation of free HDAC8. Microkinetic states over simulation time; from top i) binding rail flips around Φ of Tyr100 ii) L1:L2 salt bridge presence between Lys33 and Asp87-89 as defined in the main text iii) presence of an α-helix at residues 93-97 as defined in the main text. In agreement with the shorter simulations, the figure shows a correlation of the backbone angle Φ of Tyr100 with the interaction of the L1 and L2 loop.

In Silico Mutation of Lys33 to Glu:

Instead of performing additional long time-scale simulations to probe the effect of mutations on the sampling of states of HDAC8, we took an alternative approach. Using our already calculated trajectory for apo-HDAC8, we performed the most deleterious Lys33Glu mutation *in silico* on frames along the previously simulated trajectory using the FoldX[1] program and estimated the free-energy differences between wild-type and the mutant during the trajectory, see Figure S3.

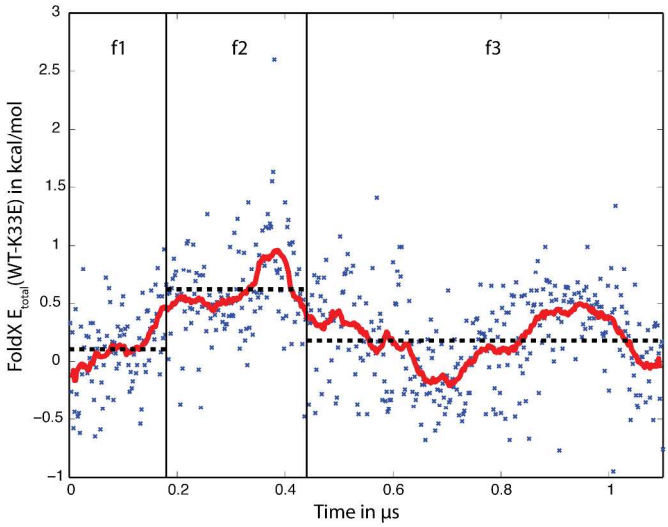

Figure S3. Estimating *in silico* the change in free energy along the MD trajectory caused by the point mutation Lys33Glu. The program FoldX[1] was used to estimate the free energy difference between the wild type HDAC8 and the Lys33Glu mutant for structures along the MD trajectory. Blue crosses show the energy differences calculated as the difference in the total energy of the mutant and the wild type over time; the red line is a smoothed representation of the data and the dashed black lines indicate the average energy difference for each of the micro-states f1-f3 shown at the top of the figure separated by black vertical lines.

Figure S3 shows that in the case of the Lys33Glu mutant the f2 state, i.e. the state in which salt-bridges form between the L1 and L2 loop, will be de-populated as the free energy difference is positive compared to the f1 and f3 state; the mean difference is 0.5 kcal/mol. This is only a trend and most likely an underestimate since FoldX optimizes the orientation of the negatively charged Glu33 side-chain in the f2 state to point away from the L2 loop. We fixed the C_β position to minimize the side-chain re-orientation by FoldX. Although this should only be taken as an approximate estimate of the change in the state energy it makes clear that using unbiased MD simulations to study the mutants would be infeasible as sampling of the f2 state would barely be achievable due to higher free energy.

(1) Guerois, R.; Nielsen, J. E.; Serrano, L. *J. Mol. Biol.* **2002**, *320*, 369–87.

HDAC8 activity derived from the MAL assay:

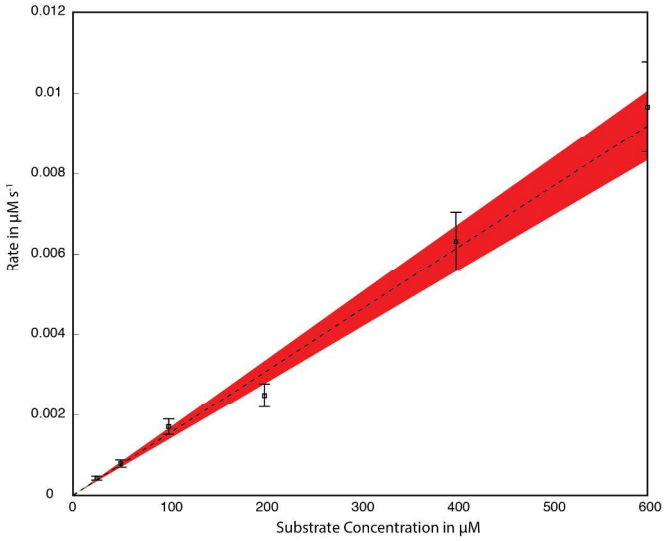

Figure S4. Determination of k_{cat}/K_M for the MAL assay. Black squares show the measured rate with error bars showing the 2σ confidence interval obtained from deriving the rate constant from the fluorescence measurement (assuming a 10% error, which was estimated from duplicate measurements). The black dashed line represents the best fit of a straight line to the data and the red area marks the 2σ confidence band from the χ^2 analysis.

The following function was fitting to the data of Figure S4:

$$v = \frac{k_{cat}\,[E_0]\,[S]}{K_M}$$

where v is the velocity or rate of the reaction, k_{cat} is the catalytic rate, K_{M} is the Michaelis-Menten constant, $[E_0]$ is the starting enzyme concentration (400 nM) and [S] is the substrate concentration. This equation is only valid for $[S] \ll K_M$, which is the case for the MAL up to concentrations where the compound becomes insoluble. The obtained rate is $k_{cat}/K_m = 38 \pm 4\ M^{-1}s^{-1}$.

NMR Activity Assay and HDAC8 Activity towards p53 Substrate:

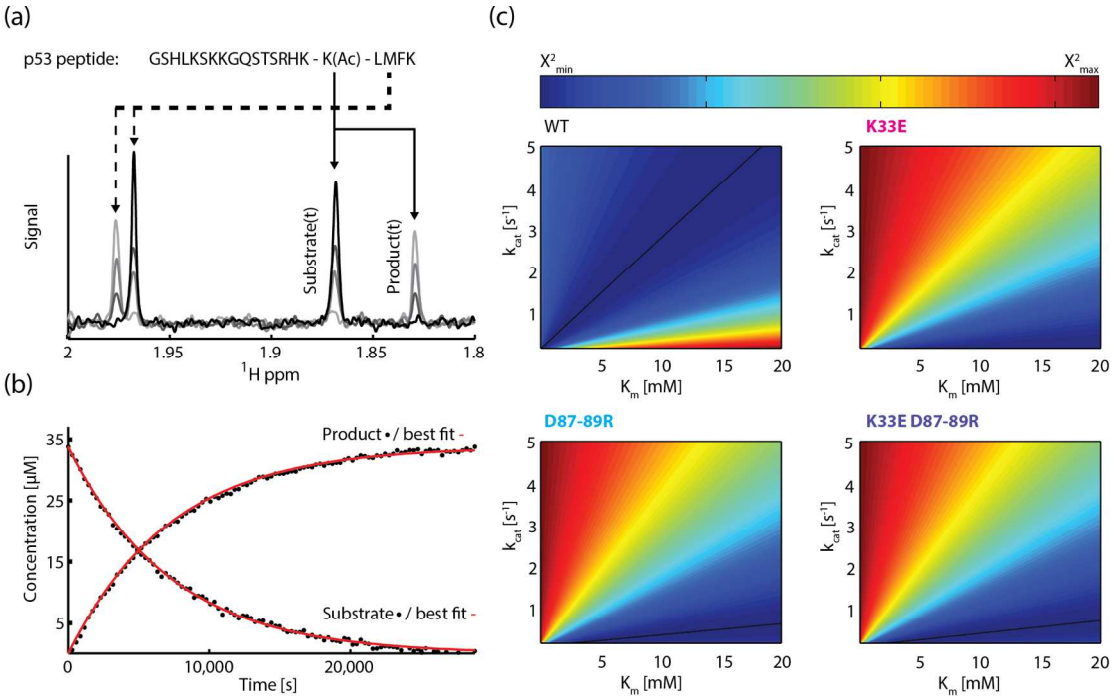

Figure S5. Real-time NMR progression HDAC8 activity assay. (**a**) 1.8 to 2 ppm chemical shift region of 1D ^{1}H NMR spectra showing the substrate and product peaks of the p53 peptide (sequence shown above the spectra) during the assay. Distinct peaks reporting on the substrate and product, respectively, are annotated and outlined by arrows, that is, the methyl group of the acetylated Lys/acetate and the methyl group the Met residue. The black spectrum shows the substrate before addition of enzyme and the time-course spectra are overlaid in fading gray. (**b**) Resulting time-course data obtained by integrating the substrate and product peak of the methionine methyl group and overlaid with the best fit obtained as described in the main text. (**c**) χ^2-error surfaces for fits of the different mutants over a k_{cat}/K_m grid. The red color corresponds to the maximum χ^2-error for each mutant and the blue color corresponds to the minimum error for each mutant.

Data Fitting: The system of ordinary differential equations (ODEs) for a Michaelis-Menten type system was numerically solved for a grid of k_{cat} and K_M values and the χ^2 calculated for each solution (see Figure S5**c**). The ODEs for a Michaelis-Menten system are:

$$\frac{d[S]}{dt} = -k_{on}\,[S][E] + k_{off}\,[ES]$$

$$\frac{d[E]}{dt} = -k_{on}\,[S][E] + (k_{off} + k_{cat})[ES]$$

$$\frac{d[ES]}{dt} = k_{on}\,[S][E] - (k_{off} + k_{cat})[ES]$$

$$\frac{d[P]}{dt} = k_{cat}[ES]$$

An initial k_{cat}/K_M was estimated along the minimal path through the χ^2 surface as a straight line. To refine the fit, perpendicular points to this line are numerically solved and χ^2 calculated, where the minimal χ^2 yields the refined k_{cat}/K_M. This is iterated three times to get the best fit for k_{cat}/K_M (red line in Figure S5**b**)). It is noted that K_m >> [S] for all experiments, which implies that it is not possible to determine k_{cat} and K_m; it is only possible to determine the ratio k_{cat}/K_M.