1   Specific Aim 2.1

SET domain protein models of SETD2 and NSD2 were built using Modeller with PDBs 4FMU and 3OOI as inputs, respectively. Proteins were placed in TIP3P water boxes of approximately 7 nm, with NaCl ions added to achieve neutrality. The ff99sb-ildn force field [3] was used. Systems were held at constant temperature (300K) and pressure (1 atmosphere) using a Langevin integrator and Monte Carlo barostat, as implemented in OpenMM 6.0.1 [2]. For each system, a total of 50-100 μs of simulations was analyzed using MDTraj [4], MSMBuilder [1], and Mixtape. Slow coordinates were selected using temporal independent component analysis (tICA) [6, 5] with randomly selected geometric features as inputs. Seven-state Markov state models were constructed by Gaussian mixture model clustering of the two slowest linear combinations of features.

References

[1] K. A. Beauchamp, G.R. Bowman, T.J. Lane, L. Maibaum, I.S. Haque, and V.S. Pande. MSMBuilder2: Modeling conformational dynamics at the picosecond to millisecond scale. *Journal of Chemical Theory and Computation*, 2011.

[2] P. Eastman, [...], and V.S. Pande. OpenMM 4: A reusable, extensible, hardware independent library for high performance molecular simulation. *Journal of Chemical Theory and Computation*, 2012.

[3] Kresten Lindorff-Larsen, Stefano Piana, Kim Palmo, Paul Maragakis, John L Klepeis, Ron O Dror, and David E Shaw. Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins: Structure, Function, and Bioinformatics*, 78(8):1950–1958, 2010.

[4] R. T. McGibbon, [...], and V. S. Pande. MDTraj: a modern, open library for the analysis of molecular dynamics trajectories. *Submitted*, 2014.

[5] Guillermo Perez-Hernandez, Fabian Paul, Toni Giorgino, Gianni De Fabritiis, and Frank Noé. Identification of slow molecular order parameters for markov model construction. *The Journal of chemical physics*, 139(1):015102, 2013.

[6] Christian R Schwantes and Vijay S Pande. Improvements in markov state model construction reveal many non-native interactions in the folding of ntl9. *Journal of chemical theory and computation*, 9(4):2000–2009, 2013.
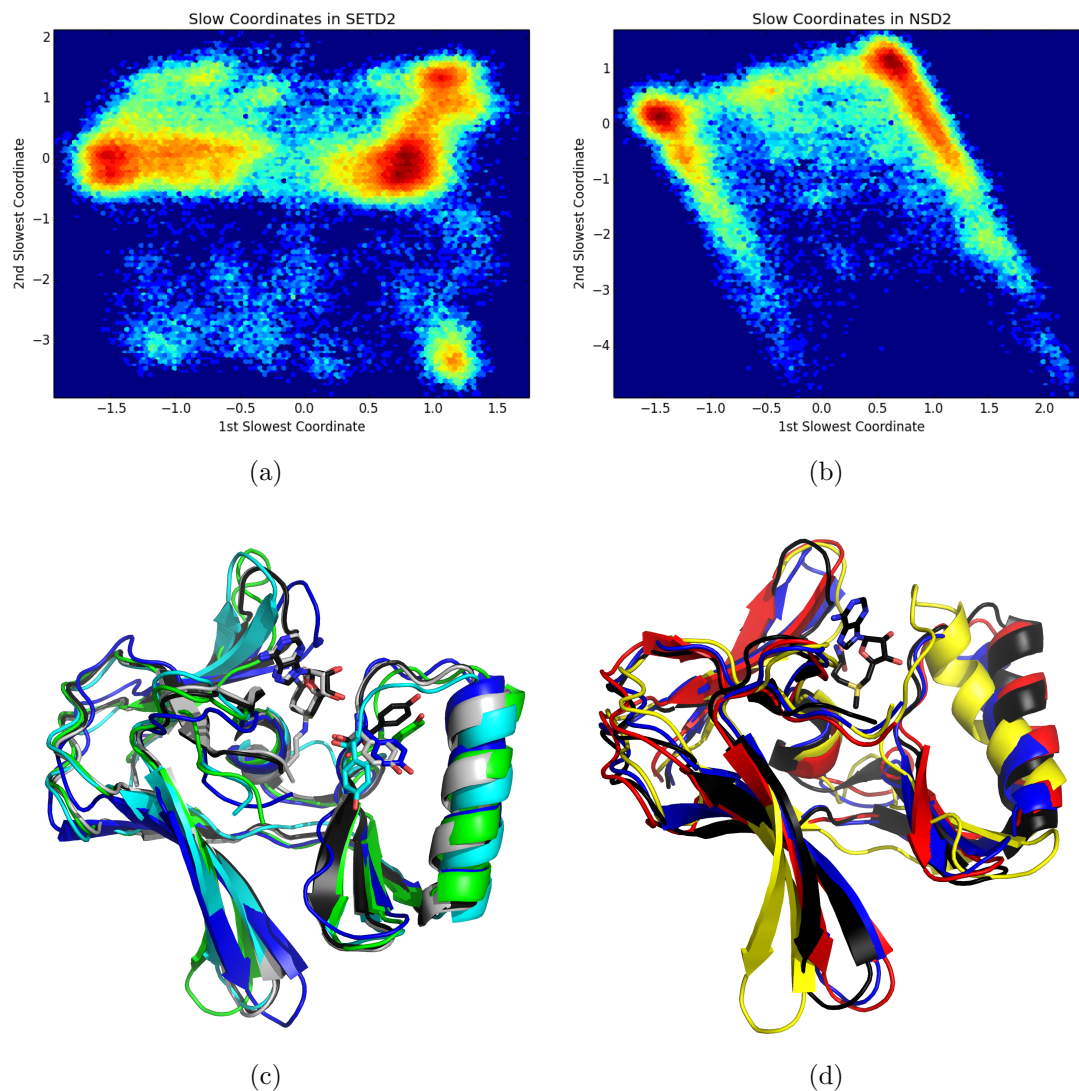
Figure 1: (a, b). Conformational landscapes—projections onto the two slowest coordinates—show multiple heterogeneous conformational states observed in simulations at the $\approx 50\mu s$ timescale. (c, d). Structural comparisons indicate conformational heterogeneity near the known SAM / sinefungin binding site. Ligand-bound crystal structures (gray, black) show similar conformational states. (c). SETD2 bound to SAH (black) adopts a TYR-up orientation (PDB:4H12), while SETD2 bound to sinefungin (gray) adopts a TYR-down orientation (PDB:4HMU); our simulation models (color) recapitulate the subtle conformational differences between these states. (d). Similar conformational heterogeneity is observed in models for NSD2 (color). Because no crystal structure is available for NSD2, NSD1 (with SAM bound) is shown for comparison (black). Simulations of SETD2 and NSD2 suggest that novel therapeutics might target and stabilize alternative near-native conformations, as has been observed in selective kinase inhibitors.