**Supplementary data**

**Zhou et al.**

**Repurposing epigenetic inhibitors to target the *Clostridioides* difficile-specific DNA adenine methyltransferase and sporulation regulator CamA**

Jujun Zhou1,3, John R. Horton1,3, Dan Yu1, Ren Ren1, Robert M. Blumenthal2, Xing Zhang1,\*, Xiaodong Cheng1,\*

1Department of Epigenetics and Molecular Carcinogenesis, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

2Department of Medical Microbiology and Immunology, and Program in Bioinformatics, The University of Toledo College of Medicine and Life Sciences, Toledo, OH 43614, USA

3 These authors contributed equally: J.Z. and J.R.H.

\* Correspondence: XZhang21@mdanderson.org or XCheng5@mdanderson.org

Email addresses:

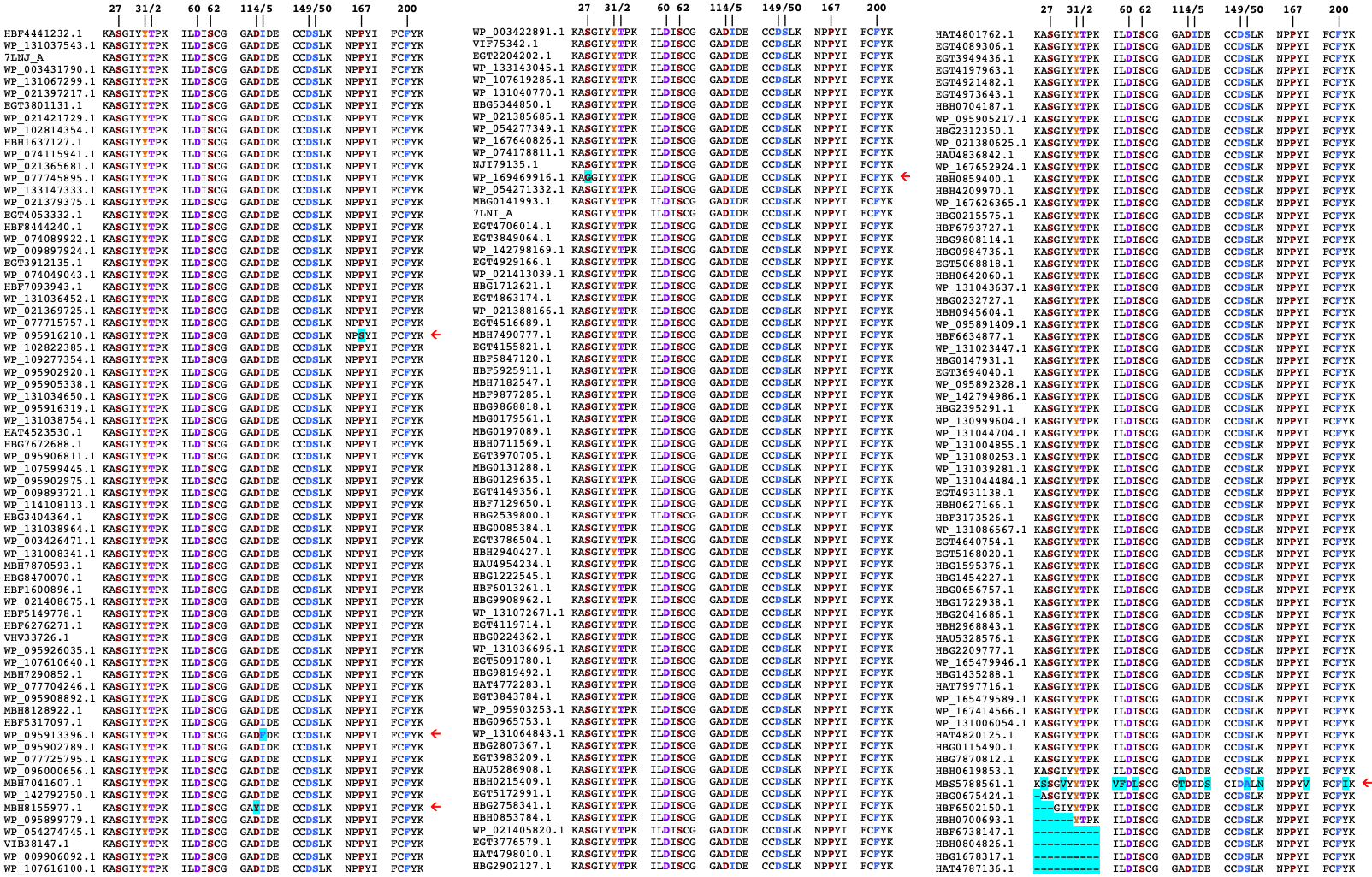
JZ (JZhou12@mdanderson.org); JRH (JRHorton@mdanderson.org);

DY (DYu6@mdanderson.org); RR (RRen1@mdanderson.org);

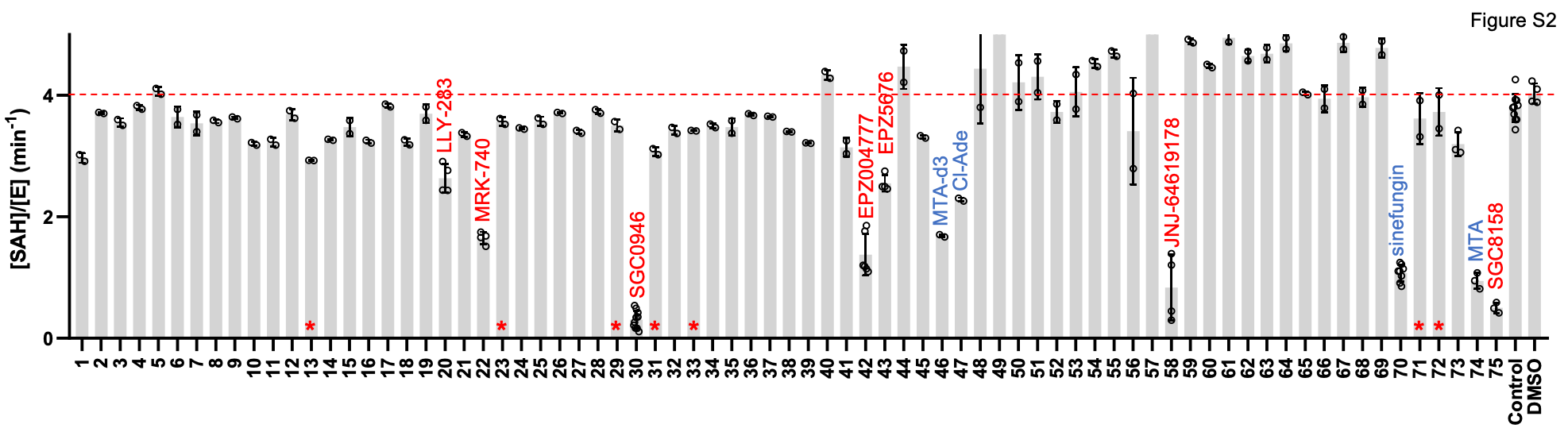
RMB (Robert.Blumenthal@utoledo.edu); XZ (XZhang21@mdanderson.org);

XC (XCheng5@mdanderson.org)

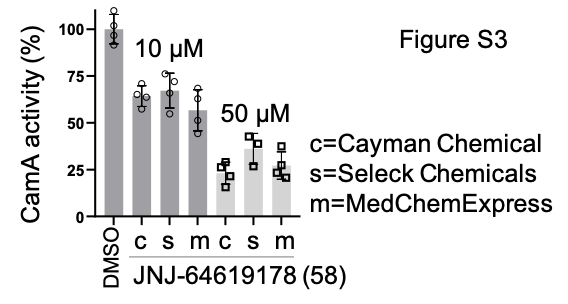
Four figures and one table



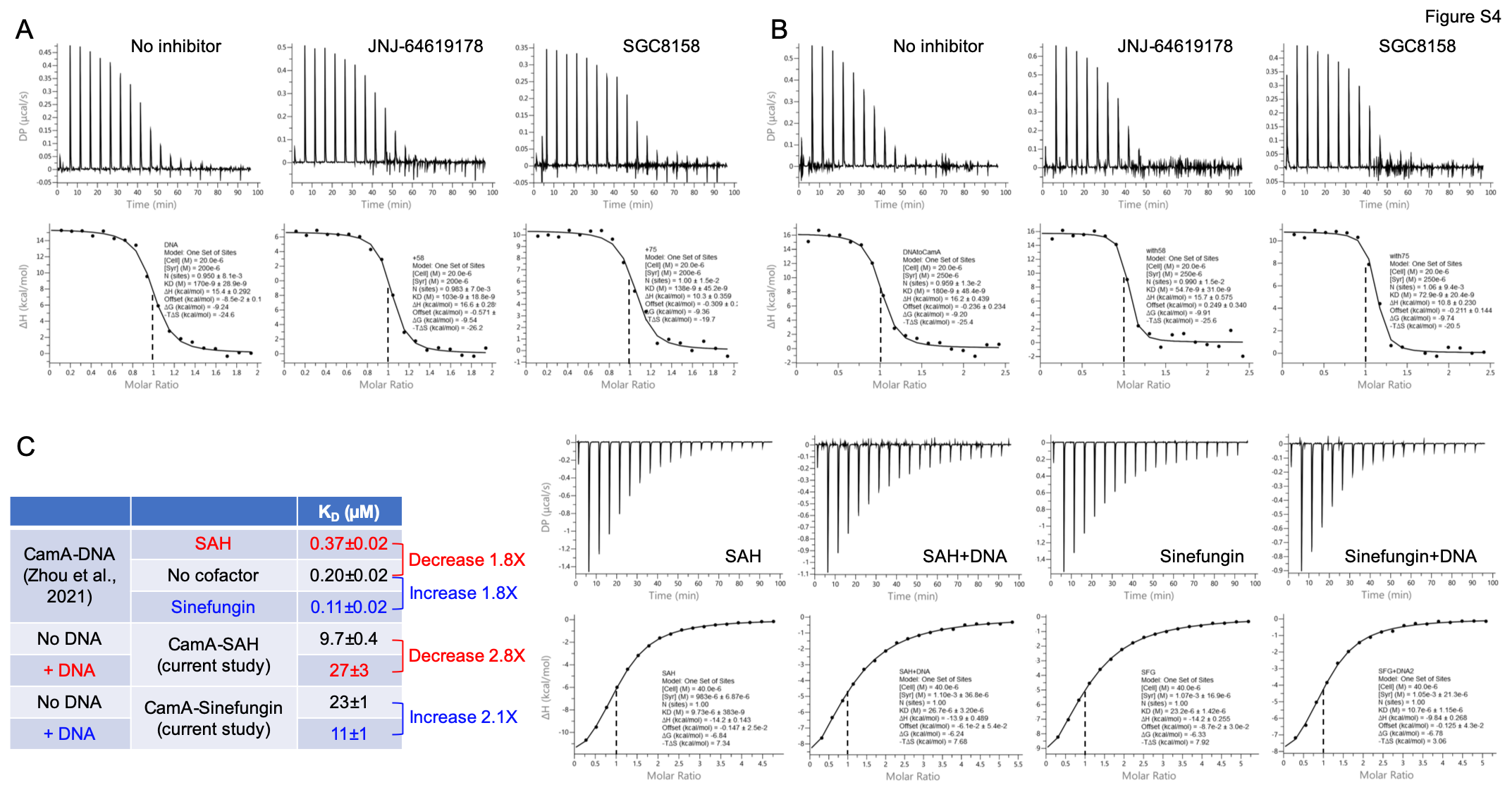
**Figure S1**. Portions of CamA orthologs that interact with SAH (PDB 7LT5; Zhou *et al.*, 2021) or sinefungin (this study), showing where substitutions have occurred (cyan highlighting, red arrows at right), plus the two flanking residues in each case. **Blue** = interacts with adenine moiety; **Purple** = interacts with aminocarboxypropyl moiety; **Red** = interacts with ribose moiety; **Orange** = interacts with both ribose and aminocarboxypropyl moieties. The 210 orthologs included here had at least 50% query coverage of the full-length CamA protein. Of these, just five had substitutions in these SAH-interacting residues. Further, of the 39 residues shown for each ortholog, four of the five orthologs with substitutions had just one, though all were in directly-interacting residues. Interestingly, ortholog MBS5788561.1 had eleven, though only one of these (S150A) is in a directly-interacting residue. Amino-terminal residues missing from shotgun sequences are also indicated, for completeness, but ignored in this analysis. In summary, this illustrates the high degree of conservation of SAM binding among CamA orthologs, suggesting that the weak SAM binding and susceptibility to competitive inhibitors is likely to be very widespread among *C. difficile* isolates.

****

**Figure S2.** Initial screen ofCamA inhibition against 75 compounds commercially available from SGC Epigenetic probe Set (https://www.thesgc.org/chemical-probes/epigenetics) purchased from Cayman Chemical Company). Blue colored ones were positive controls and red colored ones including asterisk were used for further study. The screens were conducted at compound concentrations of 50 μM. The positive controls include compound 70, sinefungin, compound 74, 5'-deoxy-5'-methylthioadenosine (MTA); compound 46, 5'-deoxy-5'-methylthioadenosine-d3; and compound 47, 5′-chloro-5′-deoxyadenosine hydrate. The negative controls include DMSO or no DMSO and no inhibitors added.

****

**Figure S3.** JNJ-64619178 from three different vendors have alike potency inhibiting CamA activity at 10 and 50 μM concentrations.



**Figure S4**. ITC measurements. (**A** and **B**) Two independent repeats of CamA-DNA binding in the absence and presence of indicated inhibitors. (**C**) CamA-cofactor binding (SAH or Sinefungin) in the absence and presence of DNA. For SAH, binding is decreased by ~2X in the presence of DNA or CamA-DNA binding is decreased in the presence of SAH. For Sinefungin, binding is increased by ~2X or CamA-DNA binding is increased in the presence of Sinefungin.

Table S1. Summary of X-ray data collection at SERCAT (22ID) beamline with wavelength=1Å

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Date Collected | 2021-02-07 | 2021-02-07 | 2021-02-07 | 2021-06-12 |
| Inhibitor | Sinefungin | SGC0946 | EPZ004777 | SGC8158 |
| PDB Code | 7RFK | 7RFL | 7RFM | 7RFN |
| Space Group | *P*212121 | *P*212121 | *P*212121 | *P*212121 |
| Cell dimensions (Å) | 82.16, 160.45, 231.30 | 81.44, 161.74, 229.85 | 81.70, 161.17, 230.61 | 81.83, 161.35, 229.91 |
| Resolution (Å) | 47.29-2.05  (2.12-2.05) | 41.87-2.38  (2.48-2.38) | 46.89-2.69  (2.79-2.69) | 47.03-2.50  (2.58-2.50) |
| a Rmerge | 0.079 (0.75) | 0.060 (0.78) | 0.088 (0.65) | 0.063 (0.92) |
| Rpim | 0.067 (0.676) | 0.062 (0.615) | 0.086 (0.818) | 0.084 (0.709) |
| CC1/2, CC | (0.395, 0.752) | (0.258,0.640) | (0.320,0.697) | (0.397,0.754) |
| b <I/σI> | 8.7 (1.1) | 10.4 (1.5) | 10.2 (1.0) | 12.3 (1.9) |
| Completeness (%) | 93.9 (84.9) | 91.6 (80.4) | 98.6 (93.9) | 99.3 (98.2) |
| Redundancy | 14.6 (6.5) | 10.8 (7.9) | 30.4 (18.1) | 12.6 (6.9) |
| Observed reflections | 2,590,541 | 1,209,001 | 2,563,747 | 1,322,596 |
| Unique reflections | 177,994 | 111,448 | 84,393 | 105,380 |
| **Refinement** |  |  |  |  |
| Resolution (Å) | 2.05 | 2.38 | 2.68 | 2.50 |
| No. reflections | 177,336 | 111,263 | 84,087 | 105,167 |
| c Rwork / d Rfree | 0.189 / 0.225 | 0.207 / 0.235 | 0.182 / 0.229 | 0.195 / 0.232 |
| No. Atoms |  |  |  |  |
| Protein | 14,014 | 13,408 | 13,319 | 13,314 |
| DNA | 1,686 | 1,686 | 1,686 | 1,704 |
| Inhibitor | 81 | 120 | 117 | 114 |
| Solvent | 1,264 | 476 | 319 | 467 |
| B Factors (Å2) |  |  |  |  |
| Protein | 44.7 | 60.6 | 65.9 | 61.9 |
| DNA | 47.5 | 67.7 | 76.2 | 72.9 |
| Inhibitor | 39.6 | 71.6 | 91.7 | 109.5 |
| Solvent | 44.0 | 50.4 | 51.4 | 52.5 |
| **R.m.s. deviations** |  |  |  |  |
| Bond lengths (Å) | 0.005 | 0.003 | 0.004 | 0.002 |
| Bond angles (˚) | 0.7 | 0.6 | 0.6 | 0.5 |

\* Values in parenthesis correspond to highest resolution shell.

a Rmerge=Σ|I-<I>|/ΣI, where I is the observed intensity and <I> is the averaged intensity from multiple observations.

b <I/σI> = averaged ratio of the intensity (I) to the error of the intensity (σI).

c Rwork=Σ|Fobs-Fcal|/Σ|Fobs|, where Fobs and Fcal are the observed and calculated structure factors, respectively.

d Rfree was calculated using a randomly chosen subset (5%) of the reflections not used in refinement.