Supplementary Information

Title: Cofactors loaded quaternary structure of Lysine-specific demethylase 5C (KDM5C) protein: Computational model

Authors: Yunhui **Peng**, Emil **Alexov**

1) pKa analysis

The residues with pKa shifts larger than 0.5 are shown in Table 1. Only two residues (Lys370 and Glu419) showed very large pKa shifts, but none of them result in change of protonation state in physiological pH range (pH=7 to 8). Human KDM5C proteins are mainly expressed in the nucleus and cytosol, where the pH ranges from 7 to 8.

Residue	pKa (unbound)	pKa(complex)	pKa shift
Lys45	10.03	11.03	1
Lys338	10.73	11.6	0.87
Lys370	10.13	12.86	2.73
Lys377	10.83	11.46	0.63
Arg390	12.18	12.69	0.51
Glu419	3.46	6.2	2.74
Lys550	10.18	10.97	0.79
Arg637	12.52	13.12	0.6
Lys711	10.47	11.1	0.63

Table 1: The pKa shifts larger than 0.5 pH units induced by formation of quaternary KDM5C structure.

2) Comparing the docking results using different servers:

In addition to the Z-dock, we tried with other two protein docking prediction servers (ClusPro and GRAMM-X) and then compared the results. The docking results comparison among the Z-dock(1), ClusPro(2) and GRAMM-X(3) serves are shown in Figure 1 and Figure 2. Figure 1 shows the ARID domain docking with the KDM5C catalytic core using different servers. The ARID domain's docking results by Z-dock and ClusPro are compared in Figure 1A. It can be seen that binding modes are almost identical. The GRAMM-X does not predict the main binding mode, but rather the secondary mode (which does not allow ARID and JmjN to be connected by 23 amino acid linker). Thus, GRAMM-X predictions are not applicable.

Figure 2 shows the comparison of the docking results for the PHD1 domain binding to the KDM5C catalytic core. Generally, the binding sites on catalytic core are consistent among different servers. The binding modes after using constrains of the linker length (marked with circle) share very similar binding modes.

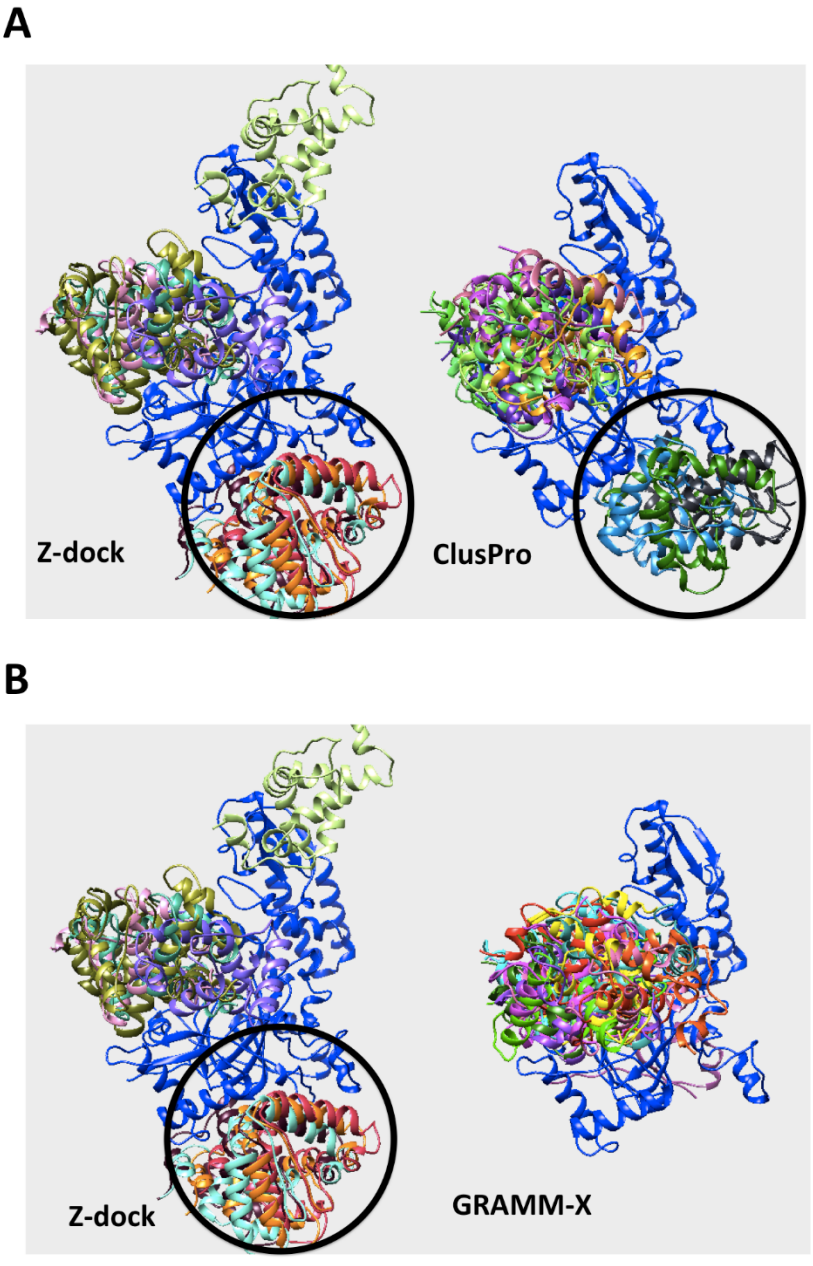

Figure 1: The ARID domain docking with the KDM5C catalytic core using different servers. The KDM5C catalytic core is marked with blue and the various ARID domain binding modes are marked with different colors. The binding modes for which ARID and JmjN can be connected by the 23 amino acids linker are circled.

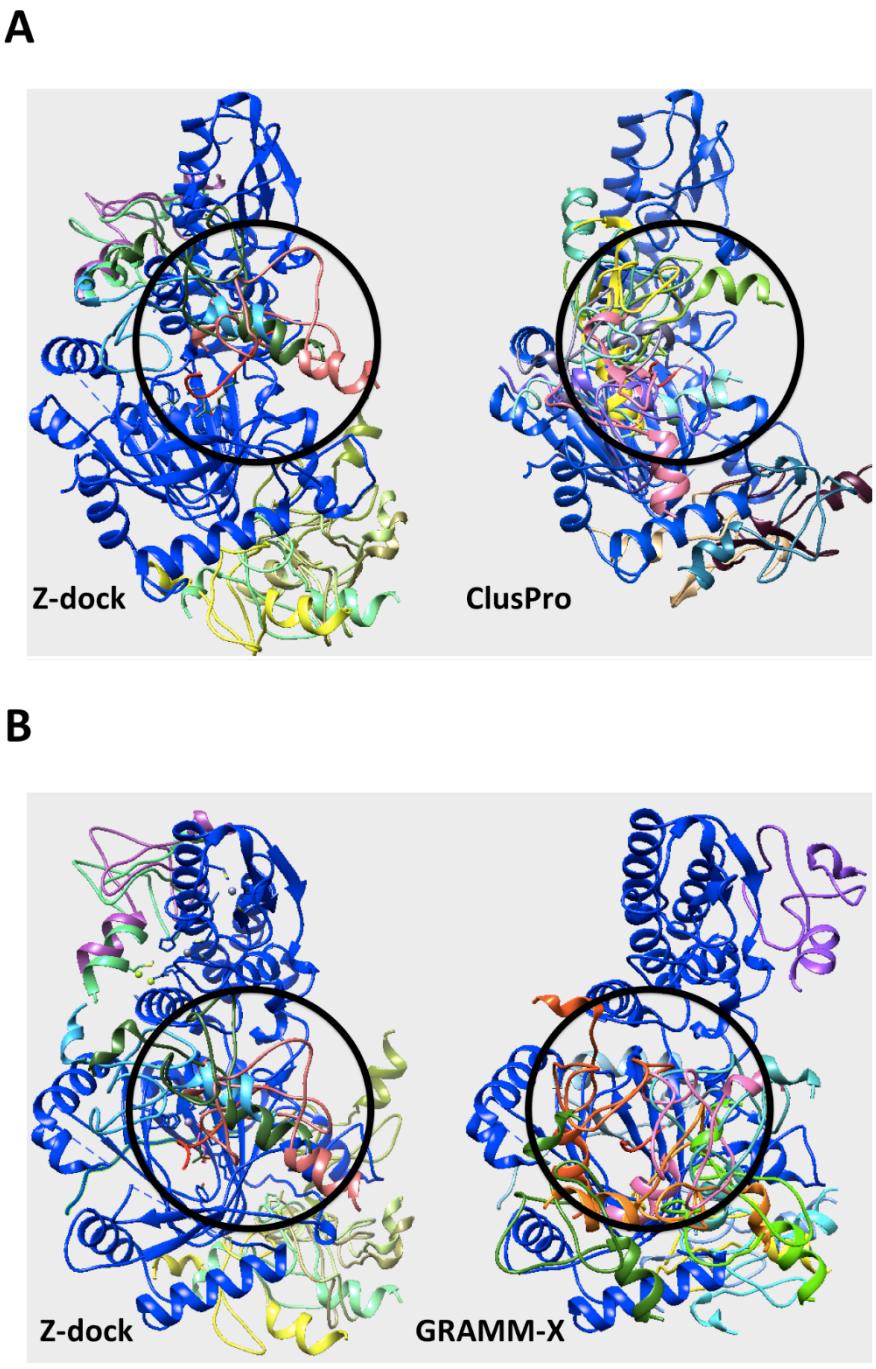

Figure 2: The PHD1 domain docking with the KDM5C catalytic core using different servers. The KDM5C catalytic core is marked with blue and the different binding modes of the PHD1 domain are marked with different colors. The binding modes, which allow PHD1 and JmjC to be connected by the 13 amino acids linker, are circled.

1. Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. Bioinformatics. 2014;30(12):1771-3.
2. Comeau SR, Gatchell DW, Vajda S, Camacho CJ. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics. 2003;20(1):45-50.
3. Tovchigrechko A, Vakser IA. GRAMM-X public web server for protein-protein docking. Nucleic Acids Res. 2006;34(Web Server issue):W310-4.