

## SUPPORTING INFORMATION

### Structure of the forkhead domain of FOXA2 bound to a complete DNA consensus site

Jun Li<sup>1,2</sup>, Ana Carolina Dantas Machado<sup>3,4</sup>, Ming Guo<sup>1</sup>, Jared M. Sagendorf<sup>3,4</sup>, Zhan Zhou<sup>1</sup>,  
Longying Jiang<sup>1</sup>, Xiaojuan Chen<sup>1,3</sup>, Daichao Wu<sup>1</sup>, Lingzhi Qu<sup>1</sup>, Zhuchu Chen<sup>1</sup>, Lin Chen<sup>1,3,\*</sup>,  
Remo Rohs<sup>3,4,\*</sup>, and Yongheng Chen<sup>1,2,5,\*</sup>

<sup>1</sup>Key Laboratory of Cancer Proteomics of Chinese Ministry of Health and Laboratory of Structural Biology, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China

<sup>2</sup>State Key Laboratory of Medical Genetics and Colleges of Life Science, Central South University, Changsha, Hunan 410008, China

<sup>3</sup>Molecular and Computational Biology Program, Departments of Biological Sciences and Chemistry, University of Southern California, Los Angeles, CA 90089, USA

<sup>4</sup>Departments of Physics & Astronomy and Computer Science, University of Southern California, Los Angeles, CA, 90089, USA

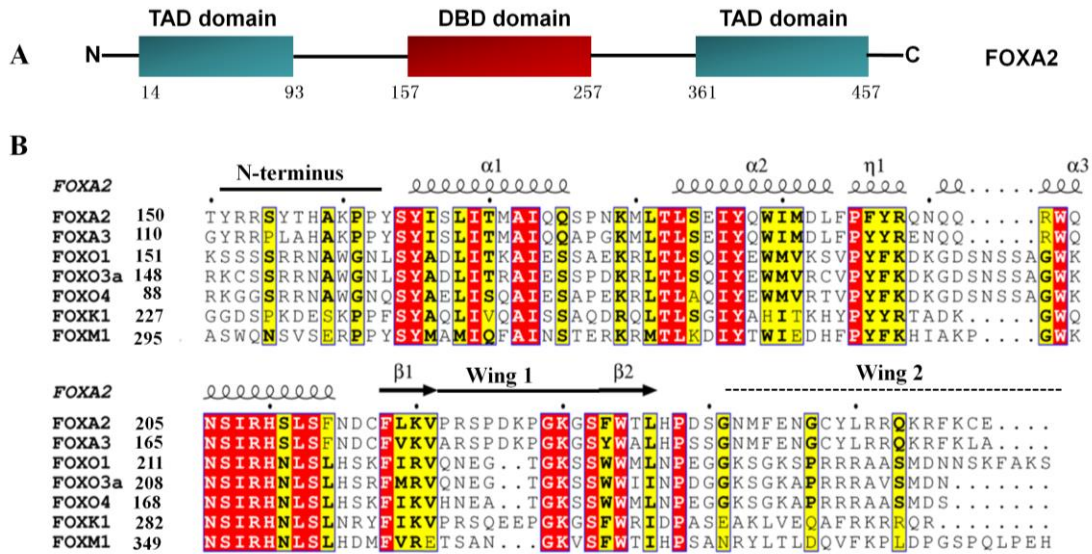
<sup>5</sup>Collaborative Innovation Center for Cancer Medicine, Guangzhou, Guangdong 510060, China

\*Corresponding authors: [yonghenc@163.com](mailto:yonghenc@163.com) (Y.C.); [rohs@usc.edu](mailto:rohs@usc.edu) (R.R.); [linchen@usc.edu](mailto:linchen@usc.edu) (L.C.)

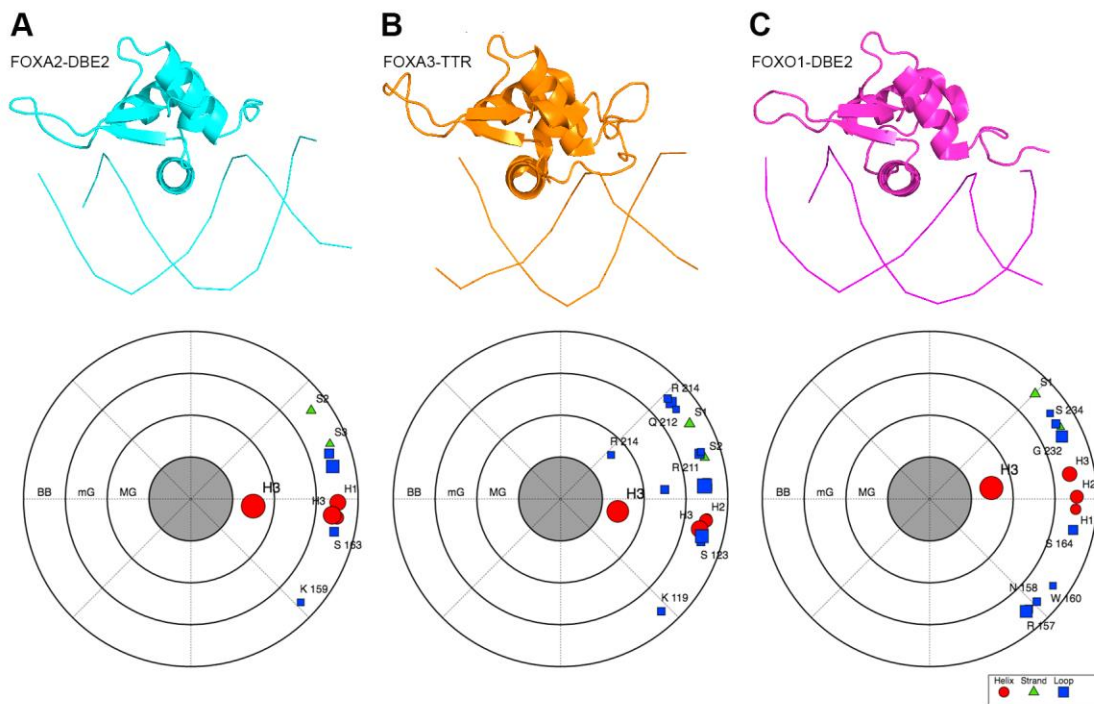
Supplementary Figure S1-----S2

Supplementary Figure S2-----S3

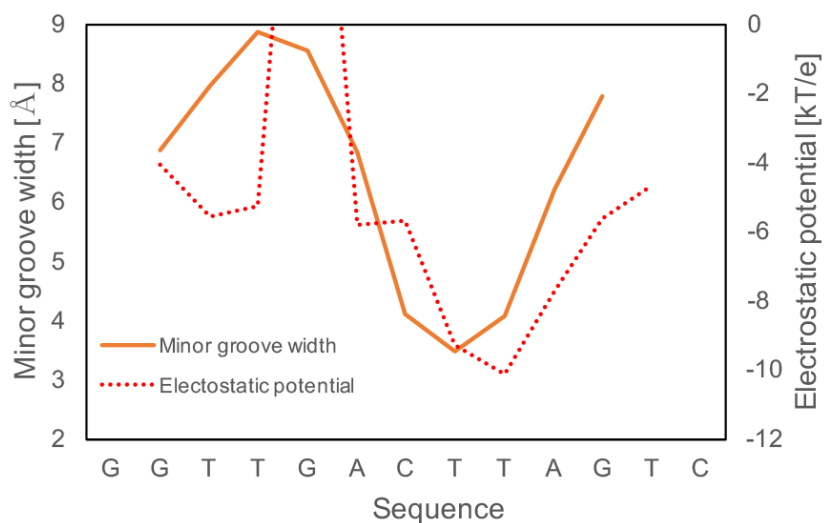
Supplementary Figure S3-----S4



**Supplementary Figure S1:** Schematic representation of FOXA2 and sequence alignment of forkhead domain with other forkhead transcription factors. (A) Domain structure of FOXA2 protein. DBD is shown in red; transcriptional activation domains (TADs) are shown in blue. (B) Sequence alignment of FOXA2 forkhead domain with other Forkhead proteins. Secondary structure elements (based on FOXA2-DBD/DBE2 complex) are indicated at the top. Identical amino acids are highlighted in red box, and similar amino acids are highlighted in yellow box.



**Supplementary Figure S2:** Comparison between overall binding signatures of FOX protein-DNA complexes. Cartoon representations (top) and schematic diagrams (bottom) are used to visualize DNA target sites bound by (A) FOXA2-DBD, (B) FOXA3-DBD, (C) FOXO1-DBD. Polar contact maps (bottom panels) are shown for backbone (outer circle), minor groove (middle circle) and major groove (inner circle) contacts. Such maps allow for visualization of different overall contacts (particularly in the loop regions) across FOX proteins when bound to their target DNA. Images were generated with DNAProDB<sup>1</sup>.



**Supplementary Figure S3:** Recognition of DNA shape and electrostatic potential. For the FOXA3-DBD/TTR complex, the narrowing of the minor groove is accompanied by an increase in negative electrostatic potential. Even though Arg210, which shows signatures of shape readout recognition, is not located at the minor groove width minimum, it is positioned in a region of increased negative electrostatic potential. Minor groove is shown as a solid orange line and electrostatic potential as a dotted red line. Electrostatic potential was calculated with DelPhi as described previously<sup>2</sup>.

#### References:

1. Sagendorf, J. M., Berman, H. M., and Rohs, R. (2017) DNAproDB: an interactive tool for structural analysis of DNA-protein complexes. *Nucleic Acids Res.* 45 (W1), W89-W97.
2. Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S., and Honig, B. (2009) The role of DNA shape in protein-DNA recognition. *Nature* 461 (7268), 1248-1253.