

Supplemental Information

Structural Insights into Histone Crotonyl-Lysine Recognition by the AF9 YEATS Domain

Qiang Zhang^{1,2,3}, Lei Zeng^{1,2,3}, ChengCheng Zhao¹, Ying Ju¹, Tsuyoshi Konuma², & Ming-Ming Zhou²

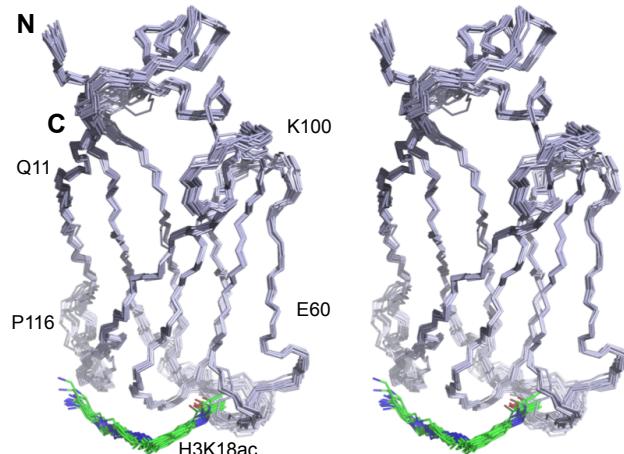
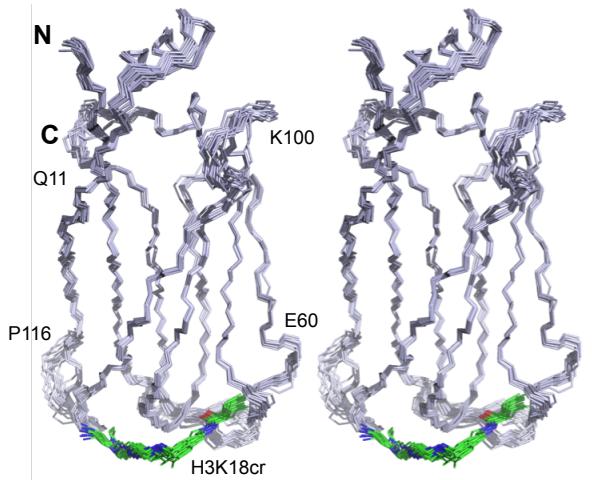
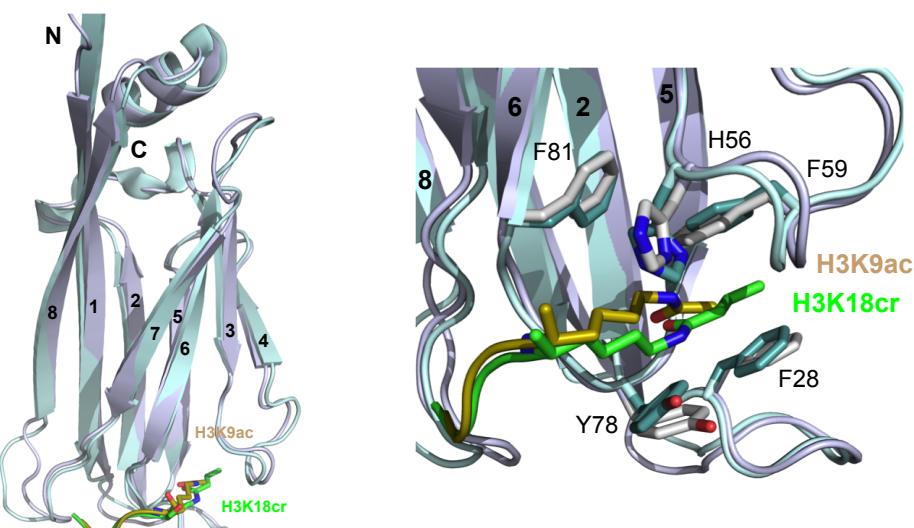
¹The First Hospital and Institute of Epigenetic Medicine, Jilin University, Changchun, 130061, China. ²Department of Structural and Chemical Biology, Icahn School of Medicine at Mount Sinai, New York, New York, 10029, USA.

³These authors contributed equally to this work.

Correspondence: ming-ming.zhou@mssm.edu (M.-M.Z.), qiang.zhang@mssm.edu (Q.Z.)

Supplemental Figure:

Figure S1. Structural analysis of the AF9 YEATS domain recognition of crotonyl-lysine in H3. Related to Figure 2.

A**B****C**

| | 1 | 20 | * | 40 | * | 60 | * | 80 |
|--------------------|--|----|---|----|---|-----|---|---------|
| AF9 (MLLT3) | MASSCA AVQVKLELGHRAQVR -KKPTVEG--- FTHDWMVFVRGP EHS-NIQHFV EKVVFH R E SFP-RP KRVCKDPPY KVEESG YAG | | | | | | | 1-80 |
| ENL (MLLT1) | MDNQCTQVRL E LGHRAQLR-KKPTTEG--- FTHDWMVFVRGP EQC-DIQHFVEKVVFWL H D S FP-KPRRVCKEPPYKVEESG YAG | | | | | | | 1-80 |
| Gas41 (YEATS4) | RVKG TIVKPIVYGNVARYFGKKREEDG -- H THQWTVVVKPYRNE-DMSAYVKKIQFKL H E S Y G -NPLRVVTKPPYEITETG WGE | | | | | | | 15-95 |
| YEATS2 | ETSRLFVKKTIVVGNVSKYI P PD K REND-QSTHKWMVYVGRSR R EPSINHV V KVWF L H P S Y K P N D L V E V R PP F H L T R G WGE | | | | | | | 200-284 |
| Yaf9 | RIKTI S VS R P I IY G NTAK M GSV K P N P A E H THL W T I F V R G P Q N E -DI S Y F I K V V F L H D T Y P -NP V R S I E A P F E L T E T G WGE | | | | | | | 8-93 |
| Taf14 | MVATVKRT I R I K T Q H I L P E V P P-VENFPV R Q W S I E V LL D E G K E-I P A T I F D K V I Y H H P T F A -NP N R T F D P P F R I E E Q G WGG | | | | | | | 1-83 |
| Sas5 | MDHSIE V TF R V K T Q Q V I P E Q N I R G N E L P LR R Q M E L L D A T G K E-VE P T I L S K C I Y H H S S F K -QP K R R L N S L P F F I E K T G WGE | | | | | | | 1-84 |
| | | | | | | | | |
| | 100 | * | * | | | 120 | | |
| AF9 (MLLT3) | F ILPI E VY F KN K EE P R K VR F D Y D L F L H L -----EGHPP V NHL R CE K L T F NN P N P T E D F R R K L K A --- | | | | | | | 81-138 |
| ENL (MLLT1) | F IMP I EV H FKN K EE P R K V C F T Y D L F L N -----EGNPP V NHL R CE K L T F NN P N P T E F R Y K L R A --- | | | | | | | 81-138 |
| Gas41 (YEATS4) | F E I II K I F I D P N-ERPV T L Y H L K L F -----QS D T N A M L G K K T V S E F Y D E M I F Q D P T A M M Q Q L L T S R --- | | | | | | | 96-159 |
| YEATS2 | F P V R V Q V H F K D S -QN K R I D I H N L K L D -----RT Y T G L Q T L G A T V V D E V L R H S G E D C I Y P Q S E --- | | | | | | | 285-345 |
| Yaf9 | F D I N I K V Y V E E A N E K V L N F Y H R L R H P Y A N P V P N D N G N E Q N T D H N S K D E V S S V Y F D E I V N E P N E E F F K I L M R --- | | | | | | | 94-171 |
| Taf14 | F P L D I S V F L E K A G - ER K I P H D L N F L Q -----ES Y E V H V I Q P L N K P L L T E E L A K G S T --- | | | | | | | 84-137 |
| Sas5 | F N L K I E C F F I G N A G K F S I E H D L T F E D D -----AY A V D Y T V D P H E F S H I N S E L K Y F D L P --- | | | | | | | 85-139 |

Figure S1. Structural analysis of the AF9 YEATS domain recognition of crotonyl-lysine in H3. Related to Figure 2.

(A) The structures of the AF9 YEATS domain in complex with H3K18cr peptide (GGKAPR-Kcr-QLATKA, residues 12-24) (left panel), or H3K18ac peptide (residues 12-24) (right panel), are depicted as stereoview of the backbone atoms (N, Ca and C') of 25 superimposed NMR structures of the complexes.

(B) Comparisons of the crystal structure of the AF9 YEATS domain/H3K9ac complex (4TMP, light cyan) and the NMR structure of AF9 YEATS domain/H3K18cr complex (light blue). RMSD of between backbones of two structures is 0.59 Å. Note that the Kcr pocket of the YEATS domain opens up slightly to accommodate the larger side chain of Kcr.

(C) Structure-based sequence alignment of the YEATS domains of human and yeast. The conserved residues at the crotonylated-lysine binding site are indicated in bold, and highlighted in red for the absolutely conserved and blue for highly conserved. The key residues that are located in the crotonyl-lysine binding site are shown in bold and by asterisk.