

## YEATS2 is a selective histone crotonylation reader

*Cell Research* (2016) 26:629–632. doi:10.1038/cr.2016.49; published online 22 April 2016

### Dear Editor,

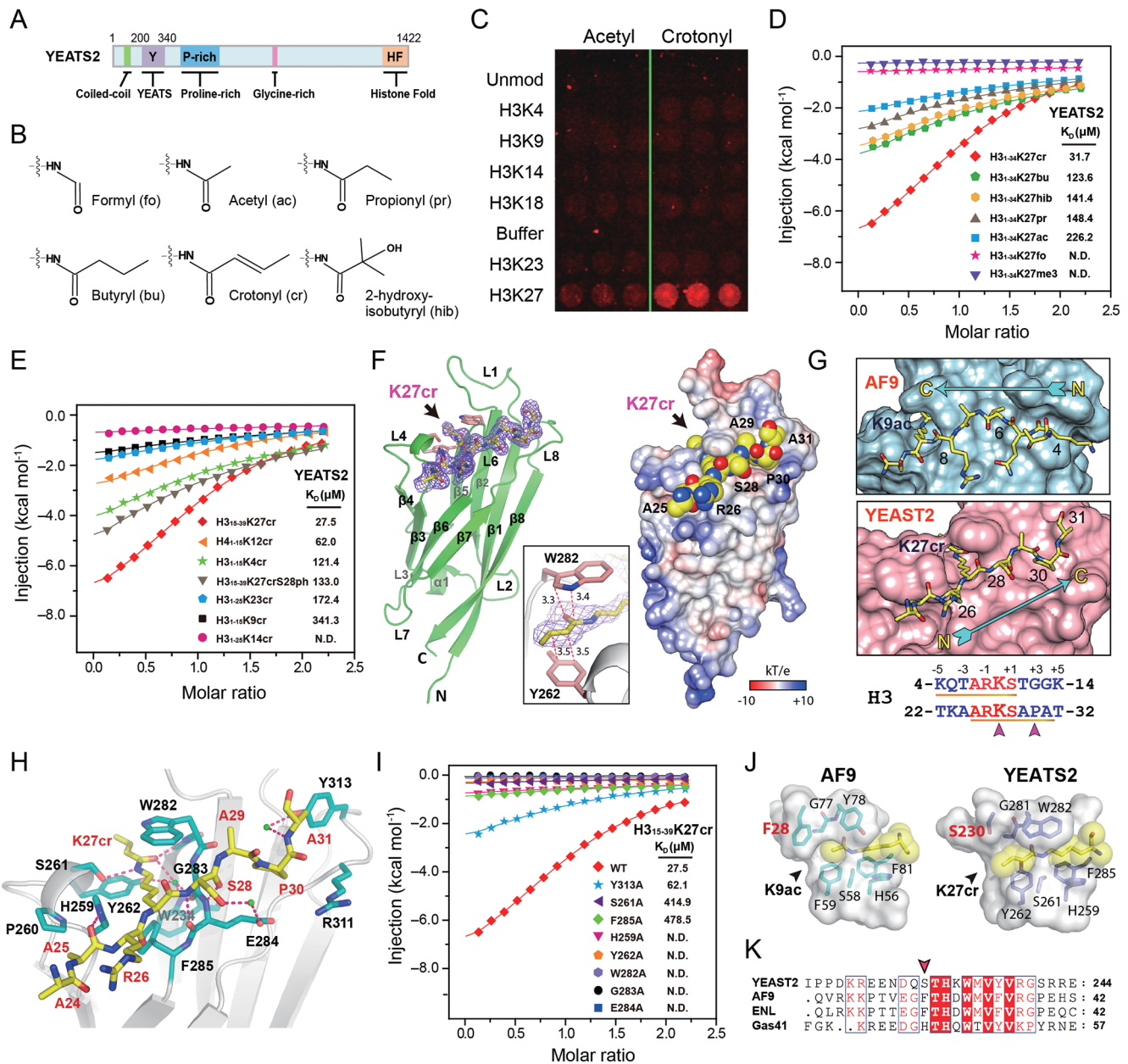
Histone lysine crotonylation (Kcr) is a recently identified non-acetyl acylation conserved from yeast to human. We have previously shown that histone Kcr marked active chromatin and was enriched at promoter and enhancer regions [1]. Different from the histone lysine acetylation (Kac), histone Kcr preferentially marks “escapee genes” during post-meiotic gene inactivation, suggesting a unique regulatory mechanism centered on Kcr independent of Kac [2]. Nevertheless, in contrast to the histone Kac readers that are well documented, selective readers for histone Kcr remain unknown. Although a few bromodomain-containing proteins (e.g., TAF1) were shown to recognize Kcr, their binding affinities are much compromised as compared to Kac binding [3].

YEATS2 is a YEATS domain-containing protein conserved from fly to human (Figure 1A). It is a stoichiometric subunit of the ADA Two A-Containing (ATAC) histone acetyltransferase complex involved in transcription regulation, stress signaling and mitotic progression [4–7]. Previously, we identified the YEATS domain as a histone Kac reader module [8]. Given the structural similarity between the acetyl and the crotonyl groups amongst other acylation types (Figure 1B), we explored the binding of recombinant YEATS2 YEATS domain to histone peptides modified by different acyl groups using a modified histone peptide array. We found that YEATS2 bound to a repertoire of acylated histone peptides with the best preference for histone H3 with crotonylation on lysine 27 (H3K27cr) in the peptide array bearing 64 reported histone acylations (Figure 1C and Supplementary information, Figure S1).

To quantify the binding property of YEATS2 towards histone peptides bearing different acylations, we performed Isothermal Titration Calorimetry (ITC) using synthetic histone peptides. We tested various acylations on H3K27, including formylation (fo), acetylation (ac), propionylation (pr), butyrylation (bu), crotonylation (cr), and 2-hydroxyisobutyrylation (hib), and used trimethylation (me3) as a negative control. Calorimetric titrations revealed that the YEATS domain of YEATS2 bound to H3<sub>1–34</sub> peptides with acylations on K27 in a preference

order of K27cr (31.7  $\mu$ M) > K27bu (123.6  $\mu$ M) > K27hib (141.4  $\mu$ M) > K27pr (148.4  $\mu$ M) > K27ac (226.2  $\mu$ M) > K27fo (N.D.) and K27me3 (N.D.) ( $K_D$  values are shown in parenthesis, N.D., not detected; Figure 1D). This suggests that YEATS2 favors acyl groups harboring longer hydrocarbon chains with the best selectivity for Kcr that discriminates against Kac by approximately 7-fold. To further validate site specificity, we synthesized peptides bearing crotonyl modifications at H3 K4, K9, K14, K23, K27, or H4 K12. Calorimetric titration revealed that the YEATS domain of YEATS2 bound to H3<sub>15–39</sub>K27cr with the strongest affinity ( $K_D$  = 27.5  $\mu$ M), and showed compromised binding activities towards H4<sub>1–15</sub>K12cr ( $K_D$  = 62.0  $\mu$ M), H3<sub>1–15</sub>K4cr ( $K_D$  = 121.4  $\mu$ M), H3<sub>1–25</sub>K23cr ( $K_D$  = 172.4  $\mu$ M) and H3<sub>1–15</sub>K9cr ( $K_D$  = 341.3  $\mu$ M). No binding was observed for H3<sub>1–25</sub>K14cr, and the adjacent S28 phosphorylation (S28ph) reduced H3<sub>15–39</sub>K27cr binding by 5-fold (Figure 1E). Collectively, these results demonstrate that the YEATS domain of YEATS2 is a selective reader for H3K27cr.

To explore the molecular basis for H3K27cr recognition by YEATS2, we determined the complex structure of the YEATS domain (aa 201–332) of human YEATS2 bound to H3<sub>24–31</sub>K27cr peptide at 2.1 Å (Supplementary information, Table S1). The overall structure of the YEATS domain adopts an immunoglobulin  $\beta$ -sandwich fold (Figure 1F, left). Based on the electron densities, we could clearly trace the H3 fragment, “A24–A25–R26–K27cr–S28–A29–P30–A31”, with the Kcr group clamped by Y262 and W282 (Figure 1F, left and middle). The H3K27cr peptide is snugly stapled into the YEATS domain with a buried solvent accessible surface area of 636 Å<sup>2</sup>. The histone peptide-binding surface of YEATS2 is less negative compared to that of AF9 [8] due to surrounding positively charged patches (Figure 1F, right), which may partly account for the relatively weak histone binding activity observed for YEATS2. Another striking difference of H3 peptide recognition between YEATS2 and AF9 is the opposite orientation of the peptides. In the complex structure of AF9–H3K9ac [8], the N-terminal motif “K4–Q5–T6–A7–R8” of H3 participates in AF9 binding (Figure 1G, top panel). However, in the complex structure of YEATS2–H3K27cr, the C-terminal motif



**Figure 1** The YEATS domain of YEATS2 is a selective reader of H3K27cr. **(A)** Domain architecture of the YEATS2 protein. **(B)** Chemical structure of different acyl groups (abbreviations correspond to data shown in **D**). **(C)** Representative peptide array results for the YEATS domain of YEATS2 (full view of the array is shown in Supplementary information, Figure S1). **(D)** ITC fitting curves of the YEATS domain of YEATS2 titrated by a series of H3<sub>1-34</sub> peptides with acylations or trimethyl-lation (me3) on K27. **(E)** ITC fitting curves of YEATS2 titrated by a series of crotonylated histone peptides. **(F)** Overall structure of YEATS2<sub>201-332</sub> bound to the H3<sub>24-31</sub>K27cr peptide in ribbon view (left) and space-filling-surface view color-coded by electrostatic potential ranging from -10 to 10 kT/e (right). YEATS2<sub>201-332</sub> (light green) is shown as ribbons, and the histone H3 peptide (yellow) is depicted as sticks. Purple mesh: Fo-Fc omit map around H3<sub>24-31</sub>K27cr peptide contoured at 2.0  $\sigma$  level. Middle, close-up view of the Kcr-sandwiching pocket; interplanar distances are labeled in the unit of angstrom. **(G)** The orientation of the H3 peptide bound to the YEATS domain. Upper, the YEATS domain of AF9 bound to the H3K9ac peptide (PDB ID: 4TMP); lower, the YEATS domain of YEATS2 bound to the H3K27cr peptide; bottom, sequence alignment around H3K9 and H3K27 sites. The consensus “ARKS” motif is highlighted in red. **(H)** Hydrogen bonding network between H3K27cr peptide and YEATS2. Hydrogen bonds are shown as pink dashes. Key residues of YEATS2 are depicted as cyan sticks and labeled black; the H3 peptide is shown as yellow sticks and labeled red. **(I)** ITC titration fitting curves of binding of YEATS2 YEATS mutants with the H3<sub>15-39</sub>K27cr peptide. **(J)** Comparison of the acyl-binding pockets of AF9 and YEATS2. Left, H3K9ac in AF9 (PDB ID: 4TMP); right, H3K27cr in YEATS2. **(K)** Sequence alignment of the acyl-binding pocket among YEATS proteins. S230 of YEATS2 is indicated with a red arrowhead.

“S28-A29-P30-A31” of H3 contributes to binding (Figure 1G, middle panel). Notably, the H3P30 inserts into a hydrophobic pocket of YEATS2 (termed P30 pocket), facilitating the positioning of H3K27cr (Supplementary information, Figure S2A). The recognition of a C-terminal sequence (+1 to +4) signature by YEATS2 partly explains its site specificity towards H3K27cr but not H3K9cr (Figure 1E) because sequences around H3K9 and H3K27 sites share a consensus “A<sub>(-2)</sub>R<sub>(-1)</sub>KS<sub>(+1)</sub>” motif, yet are divergent in flanking regions (Figure 1G, bottom). The H3<sub>24-31</sub>K27cr peptide binding is notably stabilized by a hydrogen bonding network involving residues H259, S261, Y262, W282, G283, E284, F285, and Y313 of YEATS2 (Figure 1H and Supplementary information, Figure S2B). The four-carbon crotonylamide group is planar due to  $\pi$ -electron conjugation [1]. H3K27cr adopts a trans-configuration and is notably recognized by  $\pi$ -aromatic stacking interactions within an aromatic sandwich cage formed by Y262 and W282 (Figure 1F, middle and Figure 1H). Besides, the amide group of K27cr is further recognized by relayed hydrogen bonding interactions involving residues S261 and W282 of YEATS2 (Figure 1H and Supplementary information, Figure S2B). Such acyl amide recognition accounts for the preference of YEATS2 for lysine acylation but not methylation. The ~4-fold affinity drop in the case of Kbu versus Kcr (Figure 1D) underscores the importance of  $\pi$ -aromatic interactions originated from an alkenyl feature of the crotonyl group.

To validate the crucial residues of YEATS2 for H3K27cr peptide recognition, we next performed structure-based mutagenesis studies. ITC titrations revealed that mutations of sandwich pocket residues (Y262A, W282A and F285A) and residues forming hydrogen bonds around H3K27cr (H259A and S261A) significantly weaken the interaction, highlighting their importance (Figure 1I). G283 is part of the inner wall of the K27cr-binding pocket. The complete binding loss of the G283A mutant (Figure 1I) suggests the necessity of a side chain-free G283 for proper Kcr insertion (Figure 1H). A 2.3-fold binding drop of Y313A (Figure 1I) validates a role of Y313 in H3A31 interaction (Figure 1H). The complete binding loss of E284A (Figure 1I) suggests a key role of E284 in both S28 recognition and P30 pocket formation (Figure 1H and Supplementary information, Figure S2A). Of note, the acidic feature of E284 likely explains the sensitivity of the YEATS2-H3K27cr interaction to S28ph (Figure 1E), considering the effect of charge repulsion caused by S28ph. However, this observation contrasts with the reported H3S10ph tolerance by AF9 YEATS [8].

The S230 of YEATS2 is a small side chain residue,

rendering a more open pocket around the tip of the acyl group compared to AF9 (Figure 1J). Such a feature likely accounts for the recognition of the bulky and branched Khib by YEATS2 (Figure 1D). Sequence alignment shows that YEATS2 is the only YEATS protein in human that has a small side chain residue at this position (Figure 1K), suggesting that YEATS2 is likely the only YEATS domain protein that reads Khib, a new type of histone mark that was shown to associate with active gene transcription in meiotic and post-meiotic cells [9]. In support, AF9 displayed no binding to H3<sub>1-34</sub>K27hib in the ITC assay (Supplementary information, Figure S2C).

In summary, here we identify YEATS2 as a selective histone crotonylation reader with site specificity for H3K27. Our structural study reveals an end-open aromatic sandwich pocket within YEATS2 for Kcr binding. As most Kac-binding bromodomains do not have histone Kcr-binding activity [3], selective reading of Kcr among other short-chain acyl groups by YEATS2 highlights the functional difference between the YEATS and the bromodomain families of Kac readers [9, 10]. Given the distinct genomic distributions of Kcr and Khib from Kac during cellular processes such as male germ cell differentiation [8, 9], the current study calls attention to a new regulatory mechanism that links cellular metabolism to epigenetic regulation centered on histone Kcr or Khib reading by YEATS2 and its associated ATAC complex.

#### Accession code

The atomic coordinate of YEATS2<sub>201-332</sub> bound to H3<sub>24-31</sub>K27cr has been deposited in the Protein Data Bank (PDB) under the accession number 5IQL.

#### Acknowledgments

We sincerely thank Drs Tiemei Li and Or Gozani at the Stanford University for exchanging unpublished data and scientific inputs on this manuscript. We thank the staff members at beamline BL17U of the Shanghai Synchrotron Radiation Facility for their assistance in data collection and the China National Center for Protein Sciences Beijing for providing facility support. This work was supported by the NSFC (91519304), the “973” program of MOST (2015CB910503), and the Tsinghua University Initiative Scientific Research Program to HL, and by grants from Cancer Prevention and Research Institute of Texas (RP160237) and Welch Foundation (G1719) to XS. XS is a recipient of Leukemia & Lymphoma Society Career Development Award and a R Lee Clark Fellow and Faculty Scholar of MD Anderson Cancer Center. We also acknowledge supports from the National Institute of General Medicine to CDA (GM040922), and to YZ (GM105933, DK107868, and GM115961).

Dan Zhao<sup>1,2</sup>, Haipeng Guan<sup>1</sup>, Shuai Zhao<sup>1</sup>, Wenyi Mi<sup>3</sup>, Hong Wen<sup>3</sup>, Yuanyuan Li<sup>1,4</sup>, Yingming Zhao<sup>5</sup>, C David Allis<sup>6</sup>, Xiaobing Shi<sup>3,7</sup>, Haitao Li<sup>1,8</sup>

<sup>1</sup>MOE Key Laboratory of Protein Sciences, Beijing Innovation Center for Structural Biology, Department of Basic Medical Sciences, School of Medicine, Tsinghua University, Beijing 100084, China; <sup>2</sup>College of Life Sciences, Peking University, Beijing 100871, China; <sup>3</sup>Department of Epigenetics and Molecular Carcinogenesis, Center for Cancer Epigenetics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA; <sup>4</sup>Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing 100084, China; <sup>5</sup>Ben May Department of Cancer Research, The University of Chicago, Chicago, IL 60637, USA; <sup>6</sup>Laboratory of Chromatin Biology and Epigenetics, The Rockefeller University, New York, NY 10065, USA; <sup>7</sup>Genes and Development and Molecular Carcinogenesis Graduate Programs, The University of Texas Graduate School of Biomedical Sciences, Houston, TX 77030, USA; <sup>8</sup>Collaborative Innovation Center for Biotherapy, West China Hospital, Sichuan University, Chengdu 610041, China

Correspondence: Haitao Li<sup>a</sup>, Xiaobing Shi<sup>b</sup>

<sup>a</sup>E-mail: lht@tsinghua.edu.cn

<sup>b</sup>E-mail: xbshi@mdanderson.org

## References

- 1 Tan M, Luo H, Lee S, *et al.* *Cell* 2011; **146**:1016-1028.
- 2 Montellier E, Rousseaux S, Zhao Y, *et al.* *Bioessays* 2012; **34**:187-193.
- 3 Flynn EM, Huang OW, Poy F, *et al.* *Structure* 2015; **23**:1801-1814.
- 4 Wang YL, Faiola F, Xu M, *et al.* *J Biol Chem* 2008; **283**:33808-33815.
- 5 Orpinell M, Fournier M, Riss A, *et al.* *EMBO J* 2010; **29**:2381-2394.
- 6 Suganuma T, Mushegian A, Swanson SK, *et al.* *Cell* 2010; **142**:726-736.
- 7 Krebs AR, Karmodiya K, Lindahl-Allen M, *et al.* *Mol Cell* 2011; **44**:410-423.
- 8 Li Y, Wen H, Xi Y, *et al.* *Cell* 2014; **159**:558-571.
- 9 Dai L, Peng C, Montellier E, *et al.* *Nat Chem Biol* 2014; **10**:365-370.
- 10 Chen Y, Sprung R, Tang Y, *et al.* *Mol Cell Proteomics* 2007; **6**:812-819.

(Supplementary information is linked to the online version of the paper on the *Cell Research* website.)



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Unported License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>