Published in final edited form as:

Mol Cell. 2016 April 21; 62(2): 181–193. doi:10.1016/j.molcel.2016.03.028.

Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain

Yuanyuan Li[1,2,8], **Benjamin R. Sabari**[3,8], **Tatyana Panchenko**[3,8], **Hong Wen**[5], **Dan Zhao**[1,2], **Haipeng Guan**[1,2], **Liling Wan**[3], **He Huang**[6], **Zhanyun Tang**[4], **Yingming Zhao**[6], **Robert G. Roeder**[4], **Xiaobing Shi**[5], **C. David Allis**[3,*], and **Haitao Li**[1,2,7,*]

[1]MOE Key Laboratory of Protein Sciences, Beijing Advanced Innovation Center for Structural Biology, Department of Basic Medical Sciences, School of Medicine, Tsinghua University, Beijing 100084, P.R. China

[2]Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing 100084, P.R. China

[3]Laboratory of Chromatin Biology and Epigenetics, The Rockefeller University, New York, NY 10065, USA

[4]Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10065, USA

[5]Department of Epigenetics and Molecular Carcinogenesis, Center for Cancer Epigenetics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[6]Ben May Department of Cancer Research, The University of Chicago, Chicago, IL 60637, USA

[7]Collaborative Innovation Center for Biotherapy, West China Hospital, Sichuan University, Chengdu 610041, P.R. China

SUMMARY

[*]Correspondence: alliscd@mail.rockefeller.edu, lht@tsinghua.edu.cn.
[8]Co-first author

ACCESSION NUMBERS

The Protein Data Bank accession numbers for the structure data reported in this paper are 5HJB, 5HJC, and 5HJD. The Gene Expression Omnibus (GEO) accession number for the data sets reported in this paper is GSE76539. The following link has been created to allow review of record GSE76539 while it remains in private status: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=clqvaemqjzkpxcf&acc=GSE76539

SUPPLEMENTAL INFORMATION

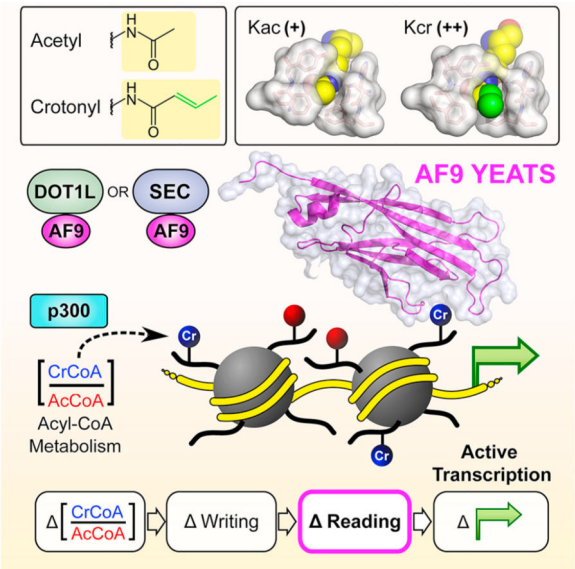
Supplemental information includes Extended Experimental Procedures, 6 figures.

AUTHOR CONTRIBUTIONS

Y.L., B.R.S. and T.P. contributed equally to this work. H.L. and C.D.A. initiated and conceived the study and wrote the manuscript; Y.L., B.R.S. and T.P. designed and performed the experiments and wrote the manuscript; H.L. and Y.L. performed structural and biophysical studies with help from D.Z. and H.G., and discovered the crotonyllysine reader activity of YEATS. B.R.S. and T.P. performed the biochemical and cellular studies under the guidance of C.D.A; H.W. and X.S. performed peptide array screening; H.H., Z.T., Y.Z., R.G.R. and X.S. provided expertise and critical feedback; L.W. generated critical reagents. The authors declare no competing financial interest.

Recognition of histone covalent modifications by chromatin-binding protein modules (“readers”) constitutes a major mechanism for epigenetic regulation, typified by bromodomains that bind acetyllysine. Non-acetyl histone lysine acylations (e.g. crotonylation, butyrylation, propionylation etc.) have been recently identified, however readers that prefer these acylations have not been characterized. Here we report that the AF9 YEATS domain displays selectively higher binding affinity for crotonyl- over acetyl-lysine. Structural studies revealed an extended aromatic sandwiching cage with crotonyl-specificity arising from π-aromatic and hydrophobic interactions between crotonyl and aromatic rings. These features are conserved among the YEATS, but not the bromodomains. Utilizing a cell-based model, we showed that AF9 co-localizes with crotonylated histone H3 and positively regulates gene expression in a YEATS domain-dependent manner. Our studies define the evolutionarily conserved YEATS domain as a family of crotonyllysine readers and specifically demonstrate that the YEATS domain of AF9 directly links histone crotonylation to active transcription.

Graphical Abstract



INTRODUCTION

Recognition of histone post-translational modifications (PTMs) by histone-binding effectors constitutes a major mechanism for epigenetic regulation (Jenuwein and Allis, 2001). A wide-range of so-called “reader” modules has been characterized for type- and site-specific readout of histone PTMs (Musselman et al., 2012; Patel and Wang, 2013; Taverna et al., 2007). Following the ground-breaking discovery of bromodomain (BrD) as an acetyllysine “reading” module (Dhalluin et al., 1999; Sanchez and Zhou, 2009), therapeutic applications of small-molecule inhibitors that block BrD-ligand interactions provide a major thrust for bromodomain study (Chi et al., 2010; Dawson et al., 2012; Marmorstein and Zhou, 2014; Yun et al., 2011). With the application of mass spectrometry-based proteomics, novel histone PTMs have been documented, including various types of non-acetyl histone lysine acylations, such as propionylation, butyrylation, crotonylation, and succinylation among

others (Huang et al., 2014; Zhao and Garcia, 2015). Alternations in cellular metabolism that, in turn, lead to shifts in the steady-state balance of non-acetyl histone acylations may function in coordinating particular transcription programs that govern cell growth and development. However, the downstream readout of histone acylations such as crotonyllysine is poorly understood in comparison to a wealth of knowledge on histone acetylation.

Like histone acetylation, histone lysine crotonylation has been detected from yeast to human and is primarily associated with active transcription (Tan et al., 2011). Crotonylation also occurs on the ε-amino group of lysine, but distinguishes itself from acetylation by its planar orientation and four-carbon length. Histone crotonyllysine (Kcr) but not acetyllysine (Kac) was found to preferentially mark “escapee genes” during post-meiotic sex inactivation in mouse testis, providing an early indication for a unique role of histone Kcr distinct from histone Kac (Montellier et al., 2012). Similarities between histone Kac and Kcr also exist, blurring distinctions. For example, Kcr and Kac sites overlap in histones and are catalyzed by p300/CBP, a well-known histone acetyltransferase (HAT). Like histone Kac, p300/CBP-mediated Kcr directly stimulates transcription in vitro and in vivo albeit to varying degrees (Sabari et al., 2015). Moreover, Sirtuin family members (e.g. SIRT1–3), well-studied histone deacetylases (HDACs), remove Kcr in a site-specific manner (Bao et al., 2014; Feldman et al., 2013). A subset of bromodomains, such as BRD9 and TAF1, can tolerate Kcr but with compromised affinity as compared to that of Kac recognition (Flynn et al., 2015). Given these findings, a clear distinction between histone Kac and histone Kcr is lacking and structural and functional consequences between the two are poorly understood.

Previously, we reported that YEATS domains constitute a novel family of histone Kac readers (Li et al., 2014). Here, through a combination of structural, biochemical/biophysical binding, and cell-based transcription activation studies, we define the evolutionarily conserved YEATS domain as a family of Kcr-favorable readers. Furthermore, we demonstrate that the YEATS domain of AF9 (AF9 YEATS) directly links histone Kcr readout to active transcription. Our work describes a new epigenetic mechanism to regulate gene activity through the establishment and readout of histone Kcr, highlighting a functional significance of non-acetyl histone acylations.

RESULTS

The YEATS Domain Has Histone Crotonyllysine Binding Activity

Recently we identified the YEATS domain as a novel Kac reader module (Li et al., 2014). Based on our crystal structural study of AF9 YEATS bound to H3K9ac, we defined a serine-lined aromatic sandwiching cage for specific readout of Kac. The long Kac side chain is sandwiched by bulky aromatic residues with the flat acetyl group snugly clamped, well-oriented in position. Interestingly, although the Kac-binding channel is long and flat, we noted a clear open space where the end of the acetyl group protrudes (Figure 1A, red arrow). Thus, we hypothesized that this opening might permit the recognition of bulkier and longer non-acetyl acylations.

To test this hypothesis, we first synthesized a series of acylated histone $H3_{1–15}K9$ peptides including formylation (fo), acetylation (ac), propionylation (pr), butyrylation (bu),

crotonylation (cr), succinylation (su) and 2-hydroxyisobutyrylation (hib) (Figure 1B), and performed isothermal titration calorimetry (ITC) using purified AF9 YEATS domain. Calorimetric titrations revealed that AF9 YEATS binds to lysine acylations in an order of K9cr (2.1 μM) > K9pr (2.7 μM) > K9bu (3.7 μM) > K9ac (5.0 μM) > K9fo (120 μM) > K9su (500 μM) > K9hib (N.D.) (K_D values shown in parenthesis, N.D., not detected, Figure 1C). Remarkably, the linear extension of the hydrocarbon chain beyond Kac resulted in 2.4-, 1.9-, 1.4-fold binding enhancement for Kcr-, Kpr- and Kbu-modified peptides, respectively. By contrast, shortened (Kfo), acidified (Ksu) or branched (Khib) K9 acylations caused a significantly reduced or complete loss of binding. Thus, our quantitative binding studies confirm that AF9 YEATS favors a subset of bulkier lysine acylations with the strongest preference for the planar, four-carbon Kcr (Figure 1C).

To determine the site preference for the AF9 YEATS-Kcr interaction, we utilized a modified histone peptide array that contains peptides bearing 64 reported histone acylations (Huang et al., 2014; Rousseaux and Khochbin, 2015; Tan et al., 2011) (For full list of peptides used see Figure S1). Consistent with the previous reports on AF9 YEATS-Kac binding (Li et al., 2014), we found that AF9 YEATS bound strongly to H3K9, H3K18 and H3K27 sites with acylation types in favor of Kcr, Kpr, Kbu and Kac, but not Ksu and Khib (Figure S1). Given the selective Kcr binding at K9, K18, K27 by AF9 YEATS in peptide array (Figure 1D), we next performed ITC to quantitatively compare the binding preference between Kcr and Kac at sites K18 and K27. We observed 2.1- and 2.7-fold binding enhancement when Kcr was introduced at $H3_{1–25}$K18 (5.7 μM Kcr *vs* 11.9 μM Kac) and $H3_{17–28}$K27 (2.7 μM Kcr *vs* 7.3 μM Kac) (Figure 1E), suggesting that Kcr can enhance the AF9 YEATS-H3 interaction as compared to acetylation regardless of acylation site differences.

Kcr has been identified in budding yeast as well as human (Tan et al., 2011). To test if Kcr preference is a conserved feature of YEATS domains across species, we recombinantly expressed and purified YEATS domains of human ENL and of yeast Yaf9 and Taf14. Calorimetric titrations using H3K9cr vs H3K9ac peptides revealed 2.5- to 4.5-fold binding enhancement for Kcr compared to Kac (Figure 1F). Collectively, these results suggest that favorable Kcr readout is a conserved function of the YEATS domain.

Structural Basis for Crotonyllysine Recognition by the AF9 YEATS Domain

To explore the molecular basis underlying Kcr readout by AF9 YEATS, we first determined the co-crystal structure of AF9 YEATS bound to $H3_{1–15}$K9cr peptide at 2.7 Å (Table 1). In the complex structure, AF9 YEATS uses the same Kac-binding aromatic sandwich cage for Kcr recognition; and the crotonyl group takes on a trans-configuration with the extended hydrocarbon chain snugly stacked against F59 aromatic ring (Figure 2A). Comparison of Kcr-bound and Kac-bound AF9 YEATS binding pocket revealed nearly identical overall pocket arrangements except for slight conformational adjustments of aromatic residues Y78 and F28 (Figure 2B, left panel). The crotonylamide group is characteristically planar due to π-electron conjugation (Figure 2B, right panel). Besides the relayed hydrogen bonding interactions conserved in both Kac and Kcr amide recognition, preferential binding to Kcr is notably contributed by π-aromatic interactions of the planar crotonylamide group with F59, Y78 aromatic rings, as evidenced by interplanar distances between 3.4–3.8 Å (Figure 2B,

right panel). Moreover, additional hydrophobic contacts introduced by hydrocarbon chain extension further stabilize Kcr-AF9 recognition. Even though a 15 amino acid long $H3_{1–15}$ peptide was used for crystallization, only H3 T3-S10 could be modelled around a surface formed by loops connecting the core β-strands of AF9 YEATS (Figure 2C). Importantly, K9cr is stapled into a pocket formed by L1, L4 and L6 loops of AF9, and the K9cr-flanking H3 residues are recognized by extensive polar or hydrophobic contacts, including a signature "H3R8-D103" hydrogen bonding pair (Figure 2C and S2A) (Li et al., 2014).

AF9 YEATS binds to $H3_{1–25}$K18cr with an affinity of 5.7 μM. We also determined the co-crystal structure of AF9 YEATS bound to $H3_{1–25}$K18cr peptide at 2.8 Å (Table 1). In the complex structure, H3 K14-L20 was modelled with K18cr recognized essentially in the same mode as that for K9cr (Figures 2D and S2B). Besides H3K18cr recognition, H3R17 at −1 position forms a charge-stabilized hydrogen bond with residue D103, recapitulating the importance of an "RK" signature motif for AF9 YEATS recognition (Li et al., 2014). Structural alignment of H3K18cr-and H3K9cr-bound AF9 complexes revealed that H3 segment "P16-R17-K18cr" overlapped well with "A7-R8-K9cr", while H3 K14-A15 and Q19-L20 displayed large discrepancy (Figures 2D and S2B). Notably, key residues constituting the AF9 YEATS binding pocket in the two complexes showed nearly identical arrangements (Figure S2C). Interaction analysis showed that hydrophobic contacts but not hydrogen bonding play a primary role for "R17-K18cr"-flanking residue recognition by AF9 (Figure S2D), which may explain why AF9 binds to H3K18cr at slightly lower affinity compared to H3K9cr.

Mutagenesis and Binding Studies

We next performed mutagenesis and ITC titration studies to verify the importance of key Kcr binding residues. As summarized in Figure 2E, alanine mutation of pocket residues F28, H56, S58, F59, G77, and Y78 resulted in 5- to 192-fold binding reduction for $H3_{1–15}$K9cr and 5- to 102-fold binding reduction for $H3_{1–25}$K18cr peptides, respectively. Strikingly, F59A led to the most dramatic binding loss, highlighting its critical role in Kcr-specific readout by providing aromatic stacking interactions. When comparing K9cr versus K9ac readout by AF9 mutants, F59A displayed more pronounced binding loss (192-fold K9cr *vs* 126-fold K9ac) while S58A (5-fold K9cr *vs* 16-fold K9ac) and Y78A (78-fold K9cr *vs* 132-fold K9ac) displayed compromised binding loss (Li et al., 2014), which highlights compensating role of F59 in dominating AF9 YEATS-Kcr interaction. D103 plays an important role in "−1" arginine recognition within the "R-Kcr" motif. In support, AF9 D103A mutation caused 36- and 10-fold binding loss for $H3_{1–15}$K9cr and $H3K_{1–25}$18cr peptides. Consistently, histone H3 R8A mutation caused a binding loss of 31fold for $H3_{1–15}$K9cr readout (Figure 2E, left panel), further validating the contribution of "H3R8-D103" hydrogen bonding pair. Sequence alignment revealed that all the pocket residues from L4 and L6 loops tested above are highly conserved among different YEATS domains (Figure 2F), consistent with the conserved Kcr-preference observed for YEATS from yeast to human.

YEATS Domain Is a Crotonyllysine Reader in the Cellular Context

To test if the AF9 YEATS-Kcr interaction occurs in the cell, we asked if we can purify crotonylated nucleosomes by using AF9 as bait. Using HeLa cell lines expressing FLAG-tagged constructs of AF9 or AF9 mutant (F59A), a point mutant that ablates YEATS-Kcr binding (Figure 2E), we generated mono-nucleosome containing nuclear extracts by MNase digestion and performed FLAG-IPs (Figures 3A and S3A). Immunoblot analysis of IP material revealed that nucleosomes marked by H3K9cr or H3K18cr co-IP with wild type AF9, but not mutant AF9 (F59A) (Figure 3B). These data demonstrate that AF9 binds nucleosomes marked by H3K9cr or H3K18cr in a YEATS-dependent manner.

We next performed nucleosome pulldowns with pre-modified nucleosomes generated *via* amber suppression. Recombinant pre-modified H3K9ac, H3K9cr, H3K18ac, and H3K18cr histones were generated by amber suppression, as previously described (Gattner et al., 2013; Kim et al., 2012) and validated by MS/MS analysis to be >90% modified (Figures 3C and S3B). These pre-modified H3 histones, in conjunction with unmodified recombinant core histones, were assembled into nucleosomes by gradient salt dialysis (Dyer et al., 2004) with a biotinylated 601 DNA template, allowing the nucleosomes to be immobilized by streptavidin (Figures S3C and S3D). The immobilized nucleosomes either unmodified, acetylated, or crotonylated (at either H3K9 or H3K18) were then incubated in HeLa nuclear extract and interacting proteins were affinity purified (Figure 3D). Western blot analysis of pulldown material, focusing on the YEATS domain proteins AF9 and ENL, revealed that nucleosomes bearing H3K9cr and H3K18cr pulled down more AF9 and ENL from nuclear extracts than nucleosomes bearing H3K9ac or H3K18ac (Figure 3D). These data corroborate the biophysical data that the YEATS domain of AF9 has a stronger preference for Kcr than Kac.

Bromodomains Lack the YEATS-like Preference for Crotonyllysine

Bromodomains constitute a major family of histone acetylation readers (Filippakopoulos et al., 2012). To explore whether bromodomains have Kcr reader activity, we expressed a variety of bromodomains and compared their binding affinity to Kac- and Kcr-containing peptides. Based on the classification of bromodomains (Filippakopoulos et al., 2012), we chose representative members from seven major BrD sub-families for the ITC titration assay. While all the bromodomains tested here exhibit reported Kac binding activity at their respective target sites, we were not able to detect cognate Kcr binding for most bromodomains (Figures 4A and S4A). Though the second bromodomain of TAF1 (abbreviated $TAF1_{BrD2}$ and in an analogous manner hereafter for other multi-bromodomain proteins) displayed an H4K5crK8cr binding activity, its binding affinity was three times weaker than the H4 acetylation counterpart (117 μM *vs* 52 μM, Figures 4A and S4A).

We next synthesized a series of histone acylation peptides bearing Kac, Kpr, Kbu, and Kcr marks (Figure 4B) at sites H4K8, H3K14, and H3K18, and performed quantitative ITC titrations for $BRD4_{BrD1}$-H4K8ac/pr/bu/cr, BAZ2A–H3K14ac/pr/bu/cr, $BRD3_{BrD2}$-H3K18ac/pr/bu/cr, as well as $AF9_{YEATS}$-H3K18ac/pr/bu/cr pairs. Interestingly, all bromodomains exhibited gradually reduced binding following chain extension from Kac to Kcr (Figures 4C and S4B), suggesting that bromodomains are primarily evolved towards the

Kac readout with restricted tolerance to Kpr, Kbu and Kcr. By contrast, AF9-YEATS exhibited enhanced binding following chain extension with tightest binding for Kcr (Figures 1C, Figure 4C, and Figure S4B). The above observations on BrD-acyllysine readout are in agreement with a recent profiling study, in which 49 bromodomains tested displayed compromised binding to non-acetyl acylations (Flynn et al., 2015). Collectively, these data suggest that bromodomains do not exhibit YEATS-like preference for non-acetyl acylation, especially Kcr.

Structural Basis for Differential Recognition of H3K18 Acylation by BRD3 and AF9

The above data suggest that BRD3 and AF9 recognize H3K18 acetyl versus crotonyl with distinct acylation type preference. To better explore the underlying structural basis, we determined the co-crystal structure of $BRD3_{BrD2}$ bound to $H3_{10–24}$K18ac peptide at 2.6 Å (Table 1). In the complex structure, H3 “A15-P16-R17-K18ac-Q19-L20-A21-T22” could be modelled with K18ac inserted into the classic Kac reader pocket formed by loops L_{ZA} and L_{BC} (Figures 4D and S4C). Interaction analysis revealed that Kac is notably stabilized by hydrophobic contacts and hydrogen bonding interactions involving an invariant asparagine N391 as well as Y348 and P333 (mediated by water) (Figures 4D and S4D). Moreover, H3K18ac flanking residues contribute multiple hydrogen bonds and hydrophobic contacts with $BRD3_{BrD2}$. Notably, histone H3R17 at −1 position forms a hydrogen bond with D394 and histone H3L20 at +2 position is inserted into a hydrophobic pocket formed by W332, P333, H395, V397 and M400, collectively accounting for H3K18 site preference (Figures 4D and S4D).

It is interesting to note that both $BRD3_{BrD2}$ and AF9 adopt a similar strategy to recognize a signature “R-K” motif around H3K18ac/cr, where the “−1R” is stabilized by an aspartate residue (D394 in BRD3 and D103 in AF9), while K18ac/cr is inserted into an adjacent reader pocket (Figures 4E and 4G). The Kac pocket of $BRD3_{BrD2}$ is formed at the center of the four-helix bundle of the bromodomain (Figures 4D and 4E); thereby natural extension of the hydrocarbon chain beyond acetylation is restricted by the dimension of the pocket as well as by a network of hydrogen bond-stabilized waters (Figure 4E). Conceivably, in order to fit into the reader pocket, bulkier acylations such as propionylation and butyrylation have to adopt a bent conformation at the cost of binding energy. This may account for about 4- and 38-fold reduced affinity for H3K18pr and H3K18bu readout by $BRD3_{BrD2}$, respectively (Figure 4C). Furthermore, given the planar feature of Kcr, H3K18cr may be too rigid and bulky to fit into the $BRD3_{BrD2}$ reader pocket, thus causing total loss of binding. In support, the modelled Kcr hydrocarbon chain directly clashes against the phenyl ring of F334, a residue highly conserved in bromodomains (Figures 4F and S4E) (Filippakopoulos et al., 2012). By contrast, the extended AF9 YEATS reader pocket is flat and end-open, ideal for interaction with acyl chains bulkier than Kac, and particularly accommodating for the rigid structure of crotonyllysine (Figure 4G).

AF9 and H3K18cr Co-Localize Genome-Wide

To assess the functional role of the AF9 YEATS-Kcr interaction, we turned to the lipopolysaccharide (LPS)-induced inflammatory response in the macrophage-like cell line RAW264.7, where we previously identified a role for p300-catalyzed H3K18cr in

inflammatory gene activation (Sabari et al., 2015). As AF9 had not been directly implicated in the inflammatory response, we first compared AF9 localization by chromatin immunoprecipitation followed by sequencing (ChIP-seq) ± LPS-stimulation for endogenous AF9 and a FLAG-tagged AF9 transgene. Both ChIP-seq data sets showed that AF9 was recruited to LPS-induced genes upon LPS stimulation (Figures 5A and S5A) and that AF9 peaks just 3' of the transcription start site (TSS) and exhibits a tail that stretches into the gene body (Figure 5A, S5A, and S5B). While ChIP-seq for both the endogenous AF9 and FLAG-AF9 demonstrated that AF9 was involved in the inflammatory response, the endogenous AF9 data suffered from high levels of background signal. Furthermore, the FLAG-AF9 protein was expressed at levels comparable to endogenous AF9 (Figure S6A). For these reasons, the remainder of the analysis was done with the higher-quality FLAG ChIP-seq data.

We next asked whether AF9 and H3K18cr co-localize across the genome. Using ChIP-seq data from LPS-stimulated cells we plotted FLAG-AF9 and H3K18cr (Sabari et al., 2015) at 4735 AF9-bound genes (Figure 5B). In this global analysis, the 3' peak of H3K18cr co-localized with the peak of AF9. Additionally, H3K18cr marked 95% of AF9-bound genes (4511 out of 4735) (Figure 5C). Similar co-localizations were observed when H3K18ac and FLAG ChIP-seq data sets were compared (Figures S5C and S5D). The co-localization of AF9 and H3K18cr can be further appreciated locally at representative genes *Ccl3*, *Slfn2*, *Nlrp3*, *Rsad2*, and *Ifit1* (Figures 5D and S5E). To determine whether there was a correlation between the levels of H3K18cr and AF9, we ranked genes marked by H3K18cr by number of read bases normalized to total reads (RPKM) ± 1kb TSS (8480 genes) and grouped these genes into quintiles Q1-Q5, with Q1 genes exhibiting the most reads for H3K18cr and Q5 the lowest (1696 genes per group). When FLAG-AF9 data was plotted for these groups, the highest average FLAG-AF9 signal was found on genes with the highest H3K18cr and vice versa (Figure 5E). In a similar analysis, RPKM values ± 1kb for genes in Q1-Q5 were counted for FLAG-AF9 and then graphed in box and whisker plots (Figure 5F). Each successive group of ranked genes exhibited significantly higher FLAG-AF9 reads, demonstrating the correlation between AF9 and H3K18cr levels at active genes (Figure 5F). The same trend was observed by plotting and counting RPKM values for H3K18cr for groups ranked by levels of FLAG-AF9 (Figures 5G and S5F). These data demonstrate the co-localization of AF9 and H3K18cr across the genome and specifically 3' of the TSS of active genes.

The AF9-Kcr Interaction Positively Regulates Gene Expression in a YEATS-dependent Manner

We have previously shown that increasing crotonyl-CoA concentrations prior to LPS stimulation (by crotonate pre-treatment) selectively increased H3K18cr levels at specific p300-targetted genes upon LPS stimulation, leading to enhanced expression (Sabari et al., 2015). Based on our findings that the YEATS domain exhibits preferential binding to Kcr, we hypothesized that this interaction and subsequent recruitment of YEATS-containing complexes could be the molecular mechanism driving the enhanced stimulation capacity of Kcr. To understand the impact of increasing H3K18cr on AF9 recruitment, we performed ChIP-seq for FLAG-AF9 from cells pre-treated with crotonate followed by LPS stimulation.

As we have previously reported, crotonate specifically enhances Kcr, and does not significantly affect Kac, because it selectively increases cellular concentrations of crotonyl-CoA, the high-energy donor of crotonyl, while leaving acetyl-CoA, the high-energy donor of acetyl, unaffected (Sabari et al., 2015). Utilizing our previously published RNA-seq data sets (Sabari et al., 2015), we generated two groups of LPS-stimulated genes (log_2(fold)>1.0; ± LPS), those that are crotonate responsive (log_2(fold)>1.0; LPS ± crotonate) and those that are crotonate unresponsive (log_2(fold)<0.01; LPS ± crotonate). Metagene analysis of FLAG-AF9 ChIP-seq data from unstimulated, LPS-stimulated, and LPS-stimulated plus crotonate pre-treatment conditions for these two sets of genes revealed that FLAG-AF9 signal was higher with crotonate treatment only for crotonate responsive genes (Figure 6A). Furthermore, by comparison of FLAG-AF9 RPKM for genes in each group, we observed a mean fold-change in AF9 of 1.58 for responsive genes compared to 0.92 for unresponsive genes (Figure 6B). These trends from metagene analysis can also be seen at a representative crotonate responsive gene (*Rsad2*) and crotonate unresponsive gene (*Ccl3*) (Figure 6C). Together these data support a role for enhanced AF9 localization at genes that have enhanced H3K18cr levels and expression.

To test whether the increase in AF9 upon increased Kcr is dependent upon the YEATS-Kcr interaction, we generated a RAW264.7 cell line expressing FLAG-AF9 (F59A) and compared this line to FLAG-AF9 (WT) by FLAG-ChIP after pre-treatment with increasing concentrations of crotonate. The two transgenes were expressed at equivalent levels (Figures S6A and S6B). Using primers designed at AF9 peaks, ChIP-qPCR analysis at a number of crotonate-responsive genes (*Rsad2*, *Il6*, *Ifit1*, *Cmpk2*, and *Gbp2*) showed that while AF9 signal increases with increasing concentrations of crotonate, the F59A mutant is unresponsive (Figures 6D and S6C). Together these data support the model that the AF9 YEATS interaction with H3K18cr is capable of enhancing AF9 recruitment. Interestingly, the F59A mutant was recruited to loci upon LPS stimulation to comparable levels as the wild type protein, suggesting that the mutant is integrated into its associated complexes and recruited by YEATS-independent means, yet it did lose the capability to “read” H3K18cr and was not further recruited upon crotonate treatment (Figures 6D and 3B).

The increase in H3K18cr by crotonate pre-treatment is associated with increased gene expression. The identification of the YEATS domain as a Kcr reader suggests that this phenomenon might be mediated through AF9. To test this hypothesis we engineered a RAW264.7 cell line with genetic knockout of AF9 by CRISPR-Cas9 (Figure S6D) and tested whether previously described genes were still crotonate responsive. By pre-treatment of parental cells with crotonate, the crotonate-responsive genes tested (*Rsad2*, *Il6*, *Ifit1*, *Cmpk2*, and *Gbp5*) exhibited between a 3- and 4-fold increase in gene expression over LPS-induction, while cells lacking AF9 demonstrated a significantly reduced response, as measured by quantitative RT-PCR (Figures 6E, S6D, and S6E). Importantly, addback of wild type AF9 but not mutant AF9 (F59A) rescued this reduction, demonstrating the gene selectivity of the phenotype and the functional importance of the YEATS-Kcr interaction in this process (Figures 6E and S6E). *Ccl3*, as a representative non-responsive gene, was unaffected by crotonate pre-treatment (Figure S6E). Both the AF9 and AF9 (F59A) transgenes were expressed at the same level (Figure S6F). Taken together these findings

establish a functional role for the YEATS-Kcr interaction in the positive regulation of gene expression.

DISCUSSION

Driven by insights gleaned from the co-crystal structure of AF9 YEATS domain in complex with H3K9ac (Li et al., 2014), we postulated that the YEATS binding pocket might prefer Kcr. This hypothesis was tested to be correct by ITC with purified components as well as by "designer nucleosome" pulldowns. The underlying molecular basis for YEATS-Kcr recognition was further revealed by co-crystal structural studies. In addition, utilizing a cell-based model of transcriptional activation wherein histone crotonylation levels can be perturbed, we showed that AF9 positively regulates gene expression in a YEATS-Kcr-dependent manner. Collectively, our work reported here demonstrates that the YEATS domain, originally identified as a novel family of Kac reader modules, exhibits an expanded acyl-binding repertoire with an evolutionarily conserved preference for Kcr over Kac. This provides the mechanism by which Kcr may function as a stronger potentiator of transcription, highlighting the functional distinction between Kcr and Kac.

A Unique Mechanism of YEATS for Preferential Crotonyllysine Readout

Kcr sets itself apart from other acylation marks by its π-electron conjugation of its crotonylamide group, thus being rigid and planar in shape (Figure 2B). Structural analysis of AF9 YEATS domain bound to H3K9cr or H3K18cr reveals that the crotonylamide plane is sandwiched by aromatic residues (Y78 from the top and F59 from the bottom), π-aromatic stacking and hydrophobic contacts, all of which contribute to preferential Kcr readout (Figure 2B). Among them, Y78 is primarily responsible for the amide group recognition, while F59 provides critical stacking contacts with the extended crotonyl hydrocarbon chain. An aromatic feature at positions 59 and 78 is highly conserved among all YEATS domains from yeast to human (F and Y at position 59, or Y and W at position 78) (Figure 2F), accounting for the conserved Kcr reader activity. Propionylation and butyrylation also enhance AF9 YEATS-H3 binding. However, due to lack of π-electron conjugation, the hydrophobic and methyl-π interactions likely dominate favorable Kpr and Kbu readout by AF9. Following Kcr insertion, AF9 F28 from the L1 loop displayed adaptive side chain flip, suggesting a role of F28 in sensing the tip part of the longer chain acylations. Thus, the occurrence of F28 may partly account for the binding loss observed for the branched Khib and acidified Ksu marks. Interestingly, residues at position 28 displayed large discrepancy across species, ranging from small side chain residue serine, as observed in YEATS2 to basic residue arginine, as seen in Sas5 (Figure 2F), suggesting that different YEATS domains may display fine-tuned acylation type sensitivity in addition to the common Kcr preference.

Comparison between the co-crystal structures of BRD3-H3K18ac and AF9-H3K18cr highlights the ways in which YEATS is distinct from bromodomains. The reader pocket of the bromodomain is generated at the center of a four-helical bundle and is side-open (Figure 4D). Kac is perpendicularly inserted into the pocket such that a natural extension of the acylation chain is blocked by the pocket dimension (Figure 4E). As supported by previous

structural studies (Flynn et al., 2015; Vollmuth and Geyer, 2010), to accommodate longer acylation chains such as propionylation and butyrylation, the extended hydrocarbon chain has to protrude out of the pocket from the side opening with compromised binding compared to Kac. However, due to the bulky and planar feature of the crotonylamide group, Kcr is too rigid to fit into the reader pockets of most bromodomains except for those that have a wider pocket, such as TAF1 (Flynn et al., 2015). By contrast, the reader pocket of YEATS is elongated and end-open, being ideal to positively sense a large repertoire of acyllysine marks with extended acylation chains (Figure 4G). Therefore, based upon our structural and functional data, we favor the general view that the YEATS-Kcr interactions studied here is fundamentally distinct from BrD-Kac interactions in BrD-containing proteins concerning both reader pocket design and histone acylation responsiveness associated with transcription control.

AF9 YEATS Links the Response of Inflammatory Genes to Histone Crotonylation

AF9 is a member of two distinct transcription-associated complexes, the Super Elongation Complex (SEC) and the Dot1L complex (DOT1LC) (Lin et al., 2010; Yokoyama et al., 2010). In this study, we found that AF9 was recruited to LPS-stimulated genes and further enhanced by crotonate pre-treatment in a YEATS-dependent manner. This crotonate enhanced AF9 localization was observed specifically at genes that have been previously shown to exhibit a crotonate-dependent increase in H3K18cr, but not H3K18ac (Sabari et al., 2015). We also observed that AF9 was recruited to sites of gene activation in a YEATS-independent manner (Figure 6D, compare 0 mM crotonate, ± LPS), likely driven by interactions with components of the DOT1L or SEC complexes, which are known to be YEATS-independent (Biswas et al., 2011; He et al., 2011; He et al., 2010; Li et al., 2014; Yokoyama et al., 2010). These observations suggest that robust and coordinated signal-induced recruitment can sometimes override the need for the YEATS-Kcr/Kac interaction, whereas the YEATS domain allows AF9 to “read” the chromatin landscape and respond to increases in histone Kcr. (Figure 6D). Furthermore, the enhanced expression observed with crotonate pre-treatment was dependent on AF9 and a fully functional YEATS domain (Figure 6E).

Our data support a role for the AF9 YEATS-Kcr interaction in the positive regulation of gene expression, but also suggest that Kcr could exert its function through currently unknown pathways. CRISPR-mediated knockout of AF9 significantly reduced the crotonate response in our model system, but did not abolish it completely (Figure 6E). This residual response could be due to trans-effects mediated by enhanced recruitment of another YEATS domain protein (ENL, GAS41, or YEAST2) or another as of yet unidentified Kcr reader(s). Histone Kcr may also function through cis-mechanisms by impacting nucleosome stability or inter-nucleosomal interactions, thereby promoting transcription in a reader-independent manner. In support of potential cis-effects, a recent study demonstrated that tetrasomes comprised of H3K122cr were less stable in solution than tetrasomes with unmodified H3 (Suzuki et al., 2015). As such, we anticipate that future studies will uncover other pathways mediated by Kcr.

We demonstrate that the YEATS domain has an expanded acyl-binding repertoire with strongest binding to Kcr. This Kcr preference prompted us to focus on the functional role of the YEATS-Kcr interaction within the context of gene activation. The enhanced affinity for Kpr or Kbu as compared to Kac for the YEATS domain is likely functional and we anticipate that perturbations to these histone modifications will impact YEATS domain localization proportionally to the measured binding affinities. The preferential binding of YEATS to Kcr as compared to Kac, reported here, makes YEATS unique among acetyllysine readers, so far tested, demonstrating that specific protein folds have evolved to functionally exploit the diversity of lysine acylations. Our work supports the general view that the wide range of lysine acylations being uncovered in histone proteins function to expand the cell's repertoire of transcriptional responses through the selective engagement of acyl-specific reader proteins.

EXPERIMENTAL PROCEDURES

Protein Production and Crystallographic Studies

All YEATS and bromodomain samples were recombinantly expressed in *E.* coli and purified as His-tagged proteins. Crystallization was performed *via* vapor diffusion method and diffraction data were collected at Shanghai Synchrotron Radiation Facility BL17U under cryo-conditions. All structures were determined by molecular replacement and refined by the program PHENIX with iterative model building by the program COOT. Detailed data collection and refinement statistics are summarized in Table 1.

ITC Titration, Mono-Nucleosome FLAG-IP and Modified Nucleosome Pulldown

All ITC titrations were performed using a MicroCal iTC200 instrument. The titration curves were fitted according to the "One Set of Sites" binding model. A high salt titration buffer (500 mM NaCl, 25 mM Tris· HCl, pH 7.5, 2 mM β-ME) was used for AF9 YEATS in order to avoid protein aggregation. Mono-nucleosome FLAG-IP was performed using stable HeLa cell lines expressing FLAG-tagged AF9 or its mutant. Mono-nucleosome pool was prepared by MNase digestion of HeLa nuclei. Kac and Kcr histone H3 were prepared using an amber suppression, and pre-modified nucleosomes were reconstituted with a biotinylated 601 DNA. Biotin-based pulldown was performed against HeLa MNase extract and analyzed by Western blot.

Cell Line Generation, ChIP, ChIP-Seq, ChIP-PCR and RT-PCR Analyses

For expression and rescue studies in RAW264.7 mouse macrophage-like cell lines, PCDH plasmids containing 3xFLAG tagged AF9 or AF9 F59A were used. Knockout of AF9 in RAW264.7 was done by the CRISPR-Cas9 strategy. FLAG- or endogenous AF9 ChIP was performed after crotonate and LPS treatment of RAW264.7 cells, followed by deep sequencing or quantitative PCR (qPCR) analyses. qPCR analyses for both ChIP-PCR and RT-PCR were carried out using the StepOnePlus Real-Time PCR system (Applied Biosystems) with selected primer pairs of target genes.

Detailed descriptions of the materials and methods are provided in the supplemental information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank members of the Li and Allis laboratories for critical discussion and input during the course of these studies. 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. We are grateful to Jason Chin and Peter Schultz for providing key plasmid reagents for amber suppression experiments, Joseph Fernandez, Brian Dill and Henrik Molina from the Rockefeller Proteomics Core for mass spectrometry analyses, Connie Zhao from the Rockefeller Genomics Core for ChIP sequencing, and Yael David from the Muir lab at Princeton University for 601 DNA plasmid and advice. This work was supported by grants from the National Natural Science Foundation of China (91519304), the Major State Basic Research Development Program in China (2015CB910503), and the Tsinghua University Initiative Scientific Research Program to H.L. We also acknowledge supports from The Rockefeller University to C.D.A and R.G.R., and the National Institute of General Medicine to C.D.A. (GM040922), to Y.Z. (GM105933, DK107868, and GM115961), and T.P. (GM112365), from CPRIT RP160237 and Welch G1719 to X.S., and from the National Cancer Institute and the Leukemia and Lymphoma Society to R.G.R.

REFERENCES

Bao X, Wang Y, Li X, Li XM, Liu Z, Yang T, Wong CF, Zhang J, Hao Q, Li XD. Identification of ‘erasers’ for lysine crotonylated histone marks using a chemical proteomics approach. Elife. 2014; 3

Biswas D, Milne TA, Basrur V, Kim J, Elenitoba-Johnson KS, Allis CD, Roeder RG. Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. Proc Natl Acad Sci U S A. 2011; 108:15751–15756. [PubMed: 21896721]

Chi P, Allis CD, Wang GG. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. Nat Rev Cancer. 2010; 10:457–469. [PubMed: 20574448]

Dawson MA, Kouzarides T, Huntly BJ. Targeting epigenetic readers in cancer. N Engl J Med. 2012; 367:647–657. [PubMed: 22894577]

Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. Structure and ligand of a histone acetyltransferase bromodomain. Nature. 1999; 399:491–496. [PubMed: 10365964]

Dyer PN, Edayathumangalam RS, White CL, Bao Y, Chakravarthy S, Muthurajan UM, Luger K. Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol. 2004; 375:23–44. [PubMed: 14870657]

Feldman JL, Baeza J, Denu JM. Activation of the protein deacetylase SIRT6 by long-chain fatty acids and widespread deacylation by mammalian sirtuins. The Journal of biological chemistry. 2013; 288:31350–31356. [PubMed: 24052263]

Filippakopoulos P, Picaud S, Mangos M, Keates T, Lambert JP, Barsyte-Lovejoy D, Felletar I, Volkmer R, Muller S, Pawson T, et al. Histone recognition and large-scale structural analysis of the human bromodomain family. Cell. 2012; 149:214–231. [PubMed: 22464331]

Flynn EM, Huang OW, Poy F, Oppikofer M, Bellon SF, Tang Y, Cochran AG. A Subset of Human Bromodomains Recognizes Butyryllysine and Crotonyllysine Histone Peptide Modifications. Structure. 2015; 23:1801–1814. [PubMed: 26365797]

Gattner MJ, Vrabel M, Carell T. Synthesis of epsilon-N-propionyl-, epsilon-N-butyryl-, and epsilon-N-crotonyl-lysine containing histone H3 using the pyrrolysine system. Chem Commun (Camb). 2013; 49:379–381. [PubMed: 23192406]

He N, Chan CK, Sobhian B, Chou S, Xue Y, Liu M, Alber T, Benkirane M, Zhou Q. Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. Proc Natl Acad Sci U S A. 2011; 108:E636–E645. [PubMed: 21873227]

He N, Liu M, Hsu J, Xue Y, Chou S, Burlingame A, Krogan NJ, Alber T, Zhou Q. HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. Molecular cell. 2010; 38:428–438. [PubMed: 20471948]

Huang H, Sabari BR, Garcia BA, Allis CD, Zhao Y. SnapShot: histone modifications. Cell. 2014; 159:458–458. e451. [PubMed: 25303536]

Jenuwein T, Allis CD. Translating the histone code. Science. 2001; 293:1074–1080. [PubMed: 11498575]

Kim CH, Kang M, Kim HJ, Chatterjee A, Schultz PG. Site-specific incorporation of epsilon-N-crotonyllysine into histones. Angew Chem Int Ed Engl. 2012; 51:7246–7249. [PubMed: 22689270]

Li Y, Wen H, Xi Y, Tanaka K, Wang H, Peng D, Ren Y, Jin Q, Dent SY, Li W, et al. AF9 YEATS domain links histone acetylation to DOT1L–mediated H3K79 methylation. Cell. 2014; 159:558–571. [PubMed: 25417107]

Lin C, Smith ER, Takahashi H, Lai KC, Martin-Brown S, Florens L, Washburn MP, Conaway JW, Conaway RC, Shilatifard A. AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. Molecular cell. 2010; 37:429–437. [PubMed: 20159561]

Marmorstein R, Zhou MM. Writers and readers of histone acetylation: structure, mechanism, and inhibition. Cold Spring Harb Perspect Biol. 2014; 6:a018762. [PubMed: 24984779]

Montellier E, Rousseaux S, Zhao Y, Khochbin S. Histone crotonylation specifically marks the haploid male germ cell gene expression program: post-meiotic male-specific gene expression. Bioessays. 2012; 34:187–193. [PubMed: 22170506]

Musselman CA, Lalonde ME, Cote J, Kutateladze TG. Perceiving the epigenetic landscape through histone readers. Nat Struct Mol Biol. 2012; 19:1218–1227. [PubMed: 23211769]

Patel DJ, Wang Z. Readout of epigenetic modifications. Annu Rev Biochem. 2013; 82:81–118. [PubMed: 23642229]

Rousseaux S, Khochbin S. Histone Acylation beyond Acetylation: Terra Incognita in Chromatin Biology. Cell J. 2015; 17:1–6. [PubMed: 25870829]

Sabari BR, Tang Z, Huang H, Yong-Gonzalez V, Molina H, Kong HE, Dai L, Shimada M, Cross JR, Zhao Y, et al. Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. Molecular cell. 2015; 58:203–215. [PubMed: 25818647]

Sanchez R, Zhou MM. The role of human bromodomains in chromatin biology and gene transcription. Curr Opin Drug Discov Devel. 2009; 12:659–665.

Suzuki Y, Horikoshi N, Kato D, Kurumizaka H. Crystal structure of the nucleosome containing histone H3 with crotonylated lysine 122. Biochem Biophys Res Commun. 2015; 469:483–489. [PubMed: 26694698]

Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell. 2011; 146:1016–1028. [PubMed: 21925322]

Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol. 2007; 14:1025–1040. [PubMed: 17984965]

Vollmuth F, Geyer M. Interaction of propionylated and butyrylated histone H3 lysine marks with Brd4 bromodomains. Angew Chem Int Ed Engl. 2010; 49:6768–6772. [PubMed: 20715035]

Yokoyama A, Lin M, Naresh A, Kitabayashi I, Cleary ML. A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. Cancer cell. 2010; 17:198–212. [PubMed: 20153263]

Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. Cell Res. 2011; 21:564–578. [PubMed: 21423274]

Zhao Y, Garcia BA. Comprehensive Catalog of Currently Documented Histone Modifications. Cold Spring Harb Perspect Biol. 2015; 7:a025064. [PubMed: 26330523]

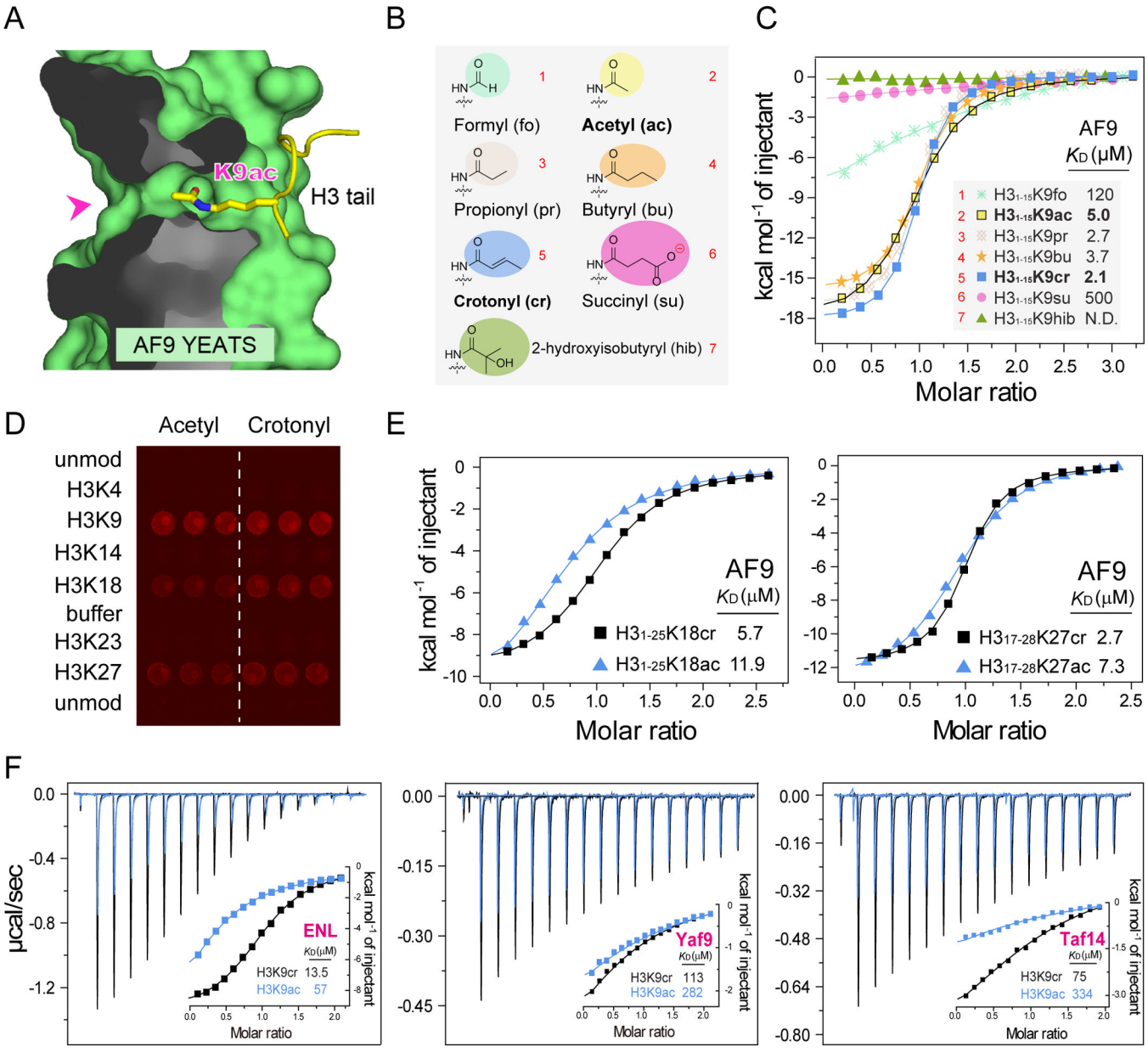


Figure 1. Identification of AF9 YEATS Domain as a Histone Crotonyllysine Reader

(A) Cut-away view of AF9 YEATS in complex with H3K9ac peptide (Li et al., 2014). AF9 YEATS is represented as green surface and histone peptide is shown in yellow. Pink arrow denotes the wide opening of Kac-reader pocket.

(B) Chemical structures of known histone lysine acylations (abbreviations and numbers correspond to data shown in (C) and are used throughout text with emphasis on acetyl- versus crotonyl-lysine).

(C) Isothermal titration calorimetry (ITC) fitting curves of wild type AF9 YEATS titrated by a series of $H3_{1–15}$ peptides containing K9 acylations shown in (B).

(D) Representative peptide array screening assay is shown in triplicate using wild type AF9 YEATS against a subset of acetylated versus crotonylated H3 peptides (for full peptide array see Figure S1).

(E) ITC fitting curves comparing Kcr and Kac binding preference at sites H3K18 (left) and H3K27 (right).

(F) ITC titration and fitting curves of H3K9ac or H3K9cr bound by YEATS domains from human ENL and yeast Yaf9, Taf14.

See also Figure S1.

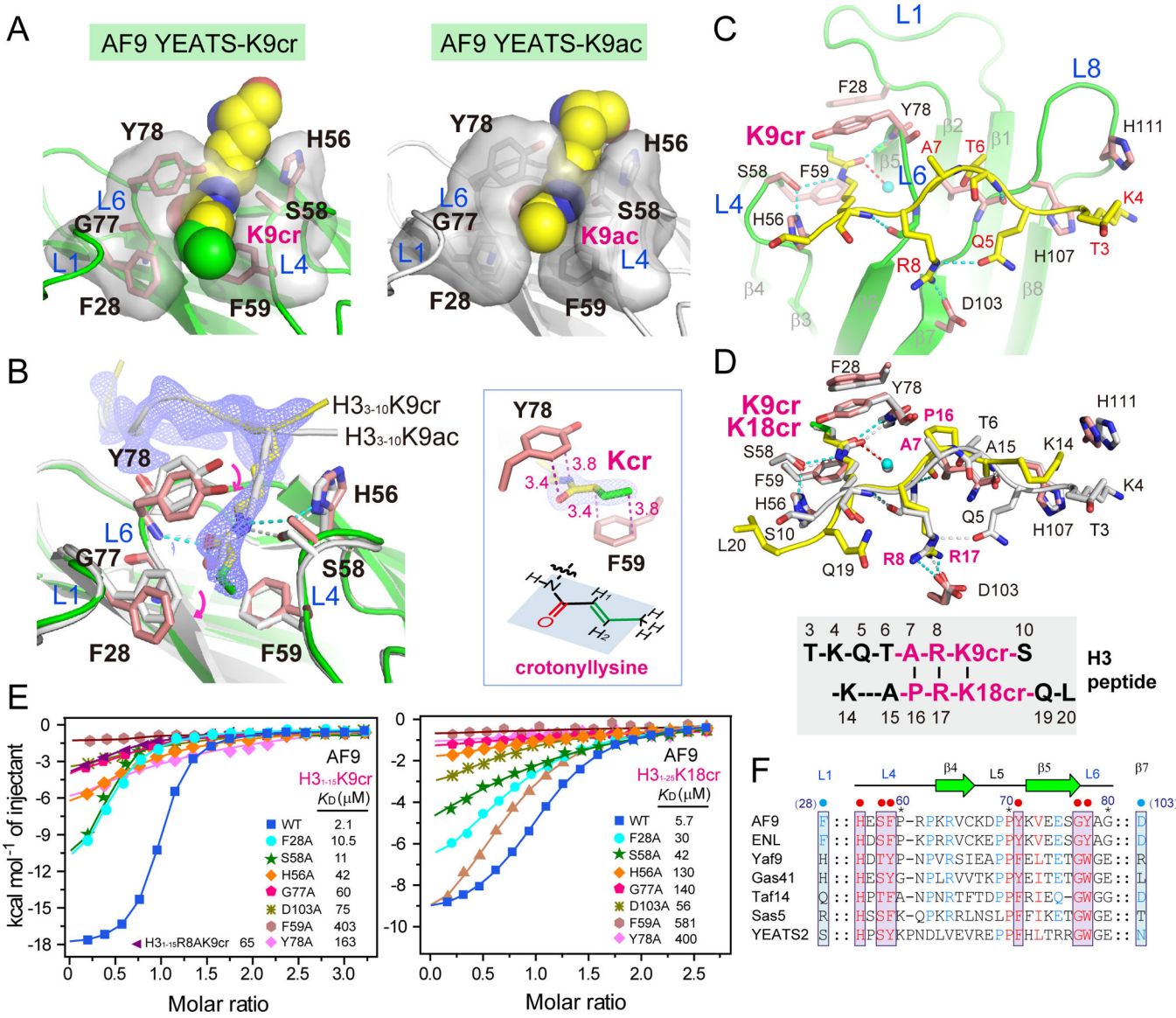


Figure 2. Molecular Basis for H3K9cr and H3K18cr Readout by AF9 YEATS domain

(A) Crystal structure of AF9 YEATS domain shows the insertion of the H3K9 containing Kcr (left) and Kac (right) into its aromatic sandwich cage. K9ac and K9cr are shown as space-filling spheres with the two additional hydrocarbon atoms of the crotonyl group shown in green.

(B) Left, detailed interaction map of the K9cr readout by AF9 YEATS shows a rotation of Y78 and F28 when compared to K9ac (pink arrows). The K9cr-bound structure (green, with key residues highlighted in light pink) is superimposed with K9ac-bound structure (white). Hydrogen bonds are shown as dashes. K9cr peptide is covered by 2Fo-Fc omit map (blue meshes) contoured at 2.5 σ. A sharpening B-factor of 75 was applied to obtain the most informative maps. Right, π-aromatic stacking between planar H3K9cr and F59 phenyl ring. Distances are denoted by magenta dashes.

(C) Overall structure of H3K9cr bound to AF9 YEATS. AF9 YEATS is shown as green ribbons with key residues highlighted in pink. Histone H3 peptide is shown in yellow.
(D) Superimposition of H3K9cr- and H3K18cr-bound complexes. H3K9cr complex is colored white as a reference and H3K18cr peptide is shown in yellow. Structure-based sequence alignment between H3K9cr and H3K18cr is shown below.
(E) Mutagenesis and ITC using mutant and wild type AF9 YEATS with H3K9cr (left) or H3K18cr (right) peptides.
(F) Sequence conservation analysis of YEATS reader pocket residues from yeast to human. See also Figure S2.

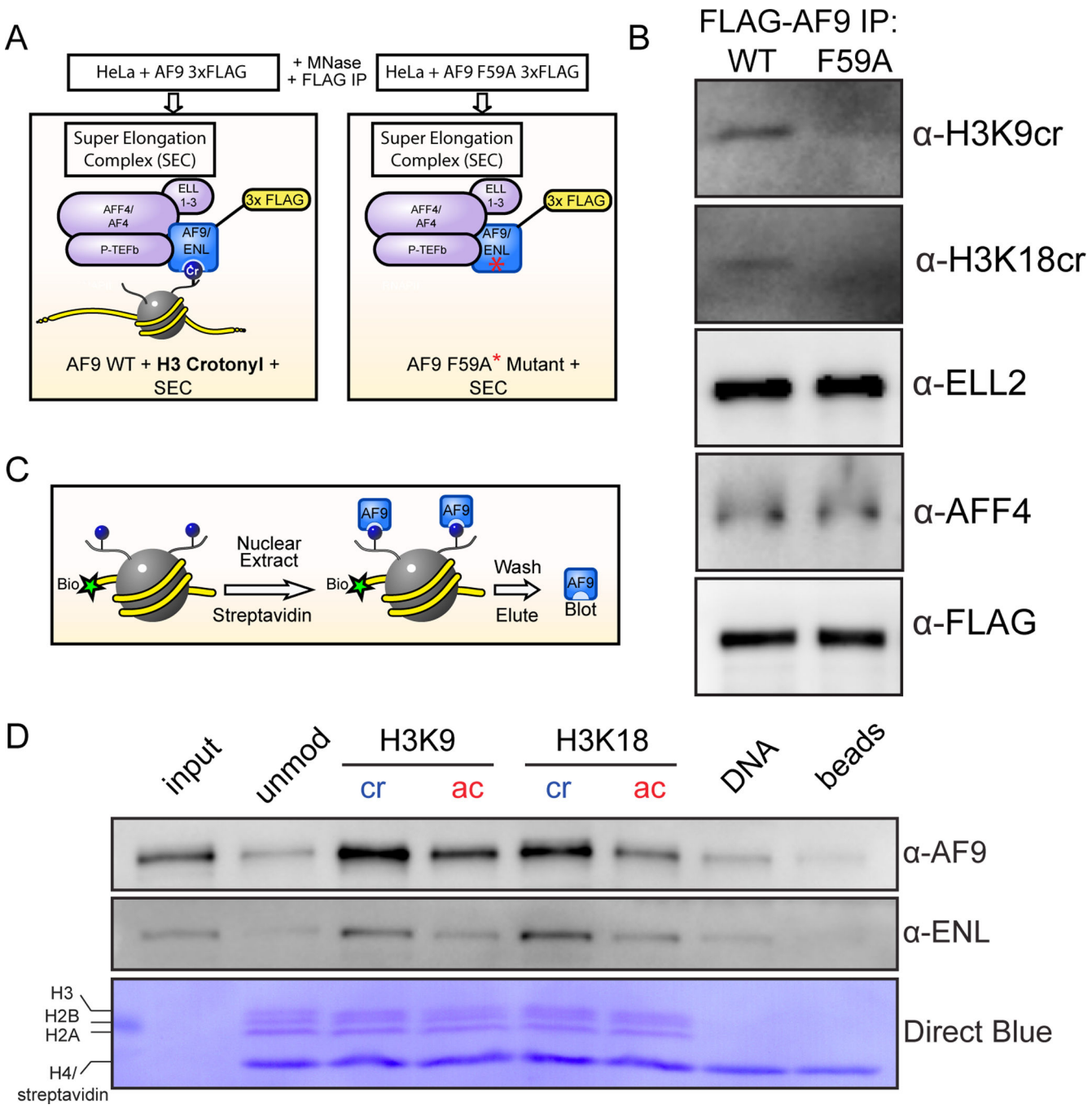


Figure 3. YEATS Domain Prefers Kcr to Kac Marked Nucleosomes

(A) Graphical illustration of AF9-FLAG MNase immuno-precipitation experiment done under native conditions. Subunits of the Super Elongation Complex (SEC) are shown with a FLAG-tagged wild type (WT) or mutant (F59A) AF9 subunits.

(B) AF9 co-immunoprecipitates nucleosomes marked by H3K9cr and H3K18cr in a YEATS-dependent manner while the interaction with the SEC components is YEATS-independent.

(C) Graphical illustration of nucleosome pulldown assay using synthetic Kcr versus Kac nucleosomes.
(D) YEATS domain proteins have a greater affinity for nucleosomes marked with Kcr than Kac. The indicated pre-modified nucleosomes generated by amber suppression were used in pull down assays. Immunoblot analysis with anti-AF9 and anti-ENL antibodies is shown. Direct Blue staining of the membrane demonstrates comparable input material. The streptavidin monomer co-migrates with histone H4 at this resolution.
See also Figure S3.

A

Bromo protein	Kcr-binding K_D(μM)	Kac-binding K_D(μM)	Bromo family	Modified site(s)
GCN5	N.D.	100	I	H4K16
BPTF	N.D.	100	I	H4K16
$BRD4_{BrD1}$	>3,000	270	II	H4K8
$BRD3_{BrD2}$	N.D.	45	II	H3K18
p300	N.D.	70	III	H4K12
BDR9	N.D.	350	IV	H3K27
BAZ2A	N.D.	20	V	H3K14
$TAF1_{BrD2}$	117	52	VII	H4K5K8
$PB1_{BrD2}$	N.D.	330	VIII	H3K14

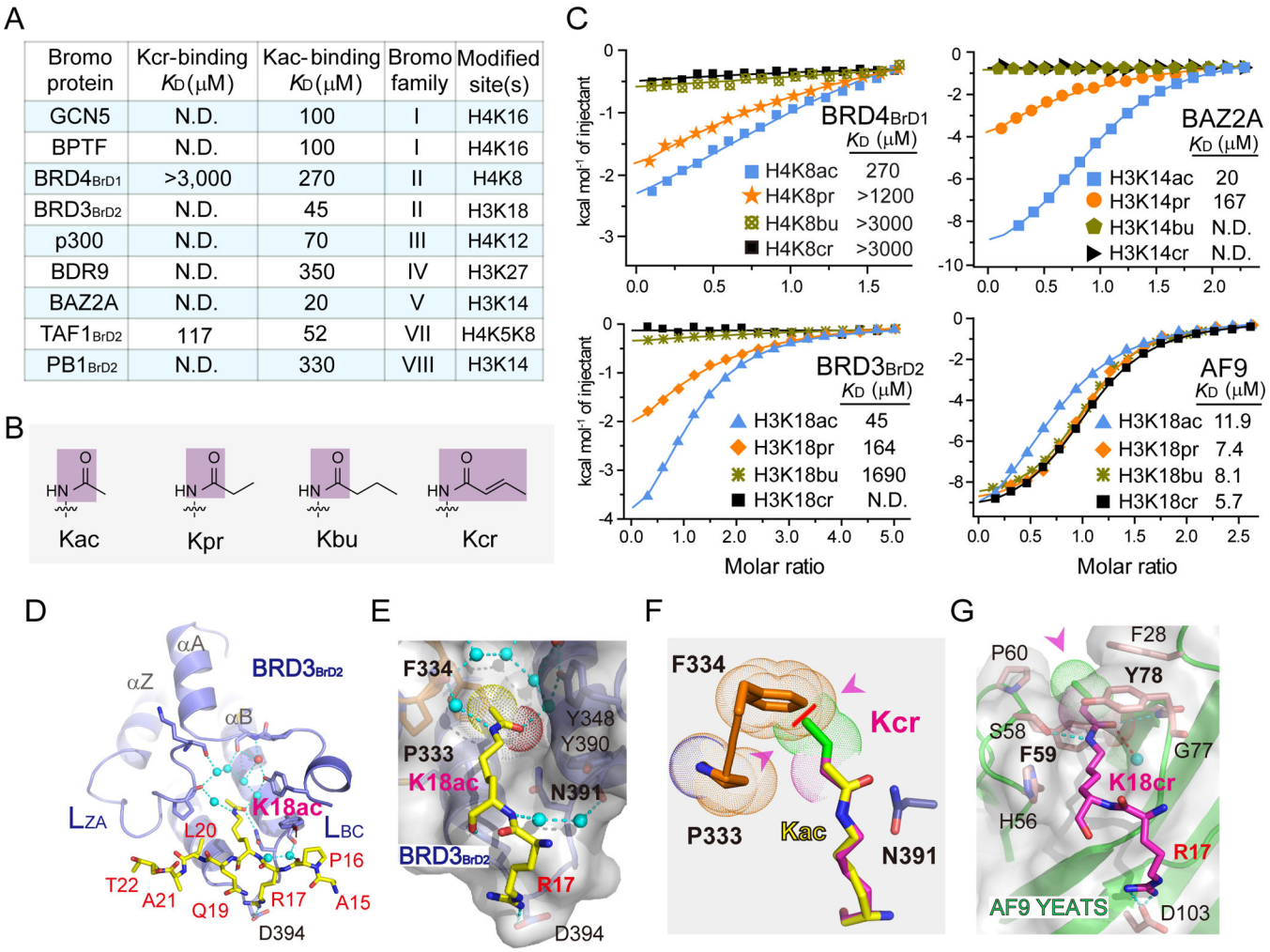


Figure 4. Bromodomains Lack Crotonyllysine Preference

(A) Summary of Kcr- and Kac-binding affinities by selected bromodomains from seven phylogenetic families (Filippakopoulos et al., 2012). Cognate modification sites are listed in the last column of the table.

(B) Comparison of chemical structures of lysine acetylation (Kac), propionylation (Kpr), butyrylation (Kbu) and crotonylation (Kcr). The planar part due to π-conjugation is shown in purple.

(C) ITC fitting curves comparing Kac, Kpr, Kbu and Kcr binding affinities of H4K8 readout by the first bromodomain of BRD4 ($BRD4_{BrD1}$), H3K14 readout by BAZ2A bromodomain, H3K18 readout by the second bromodomain of BRD3 ($BRD3_{BrD2}$) and AF9 YEATS.

(D) Histone H3K18ac binding by $BRD3_{BrD2}$. $BRD3_{BrD2}$ is shown as ribbon. H3K18ac peptide is shown in yellow. Water molecules are shown in cyan.

(E) Recognition of H3 “R17-K18ac” by $BRD3_{BrD2}$. $BRD3_{BrD2}$ is shown as surface view. Note the spatial restraints around the K18 acetylamide group caused by the side-open pocket.

(F) Steric clash between F334 of $BRD3_{BrD2}$ and a modelled K18cr. The experimental K18ac group is shown in yellow and overlaid for reference. The extended hydrocarbon group of crotonylation is colored green. Red disk indicates steric clash.
(G) Recognition of H3 “R17-K18cr” by AF9 YEATS. Note the position of crotonylamide group in the extended and end-open pocket.
See also Figure S4.

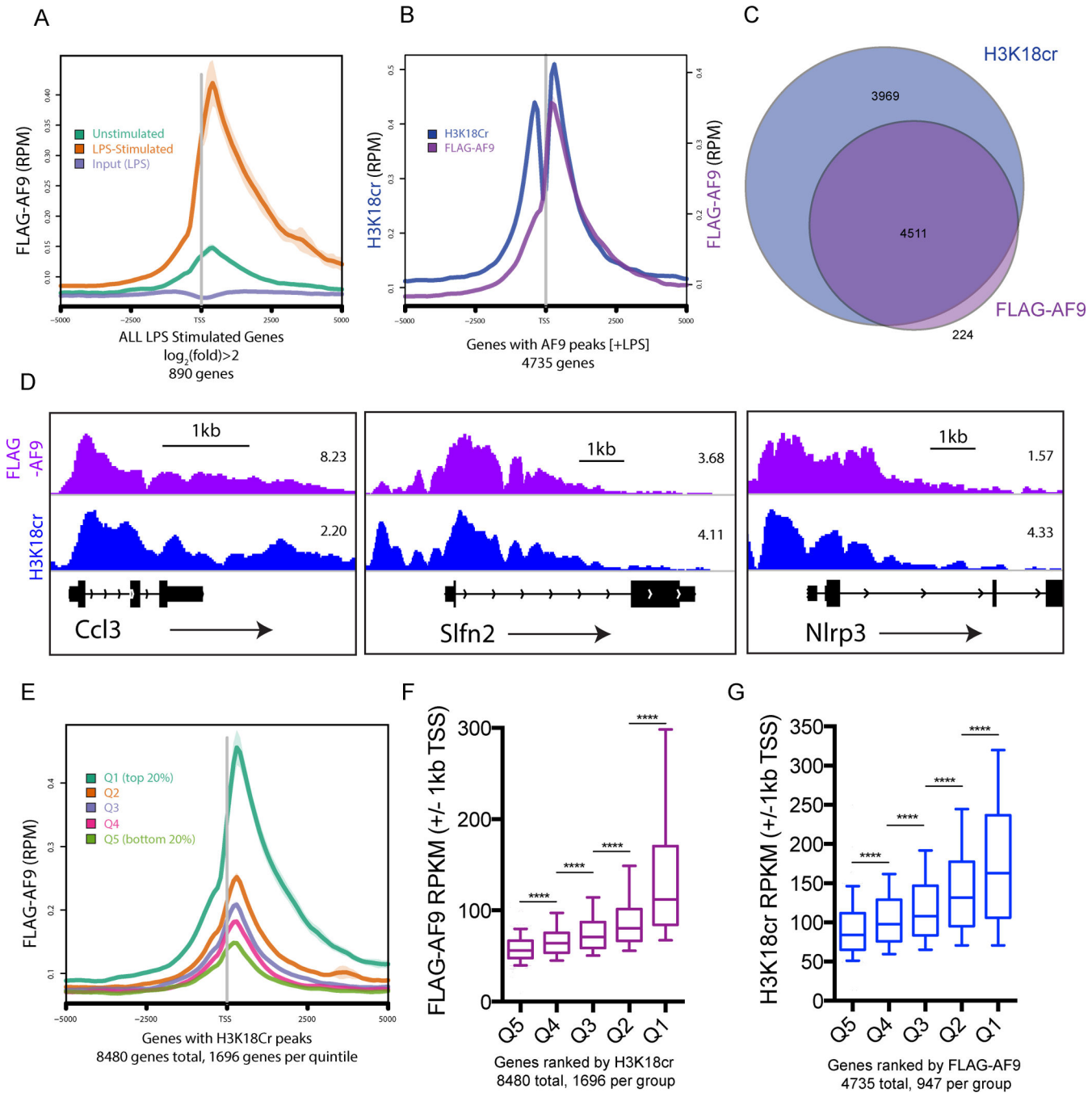


Figure 5. AF9 is Recruited to LPS Stimulated Genes and Correlates with H3K18cr

(A) AF9 is recruited to LPS stimulated genes. Average profile of FLAG-AF9 ChIP-seq data from unstimulated or LPS-stimulated RAW264.7 cells plotted ± 5kb TSS of LPS-stimulated genes (log_2(fold)>2; 890 genes).

(B) H3K18cr and AF9 co-localize at genes marked by AF9. Average profile of FLAG-AF9 and H3K18cr ChIP-seq data from LPS-stimulated cells plotted ± 5kb TSS of genes occupied by AF9 (4735 genes).

(C) Venn diagram showing the overlap of AF9 occupied (purple) and H3K18cr marked (blue) genes.
(D) Genome-browser view of FLAG-AF9 (purple) and H3K18cr (blue) at three representative AF9-occupied genes.
(E-F) FLAG-AF9 levels correlate with levels of H3K18cr. (E) Average profile of FLAG-AF9 ChIP-seq data plotted over quintiles ranked by H3K18cr RPKM ± 1kb TSS. Q1 represents the group with highest H3K18cr RPKM and Q5 the lowest (1696 genes per quintile). (F) Using the same groups, FLAG-AF9 RPKM for each gene ± 1kb within a group are plotted as a box and whisker plot (10–90th percentile) for each group. ****: t-test derived p-value < 0.0001.
(G) H3K18cr levels correlate with levels of FLAG-AF9. Similar analysis as in (F), except H3K18cr RPKM for each gene ± 1kb within quintile groups ranked by FLAG-AF9 RPKM ± 1kb TSS are plotted as a box and whisker plot (10–90th percentile) for each quintile (947 genes per quintile). ****: t-test derived p-value < 0.0001.
See also Figure S5.

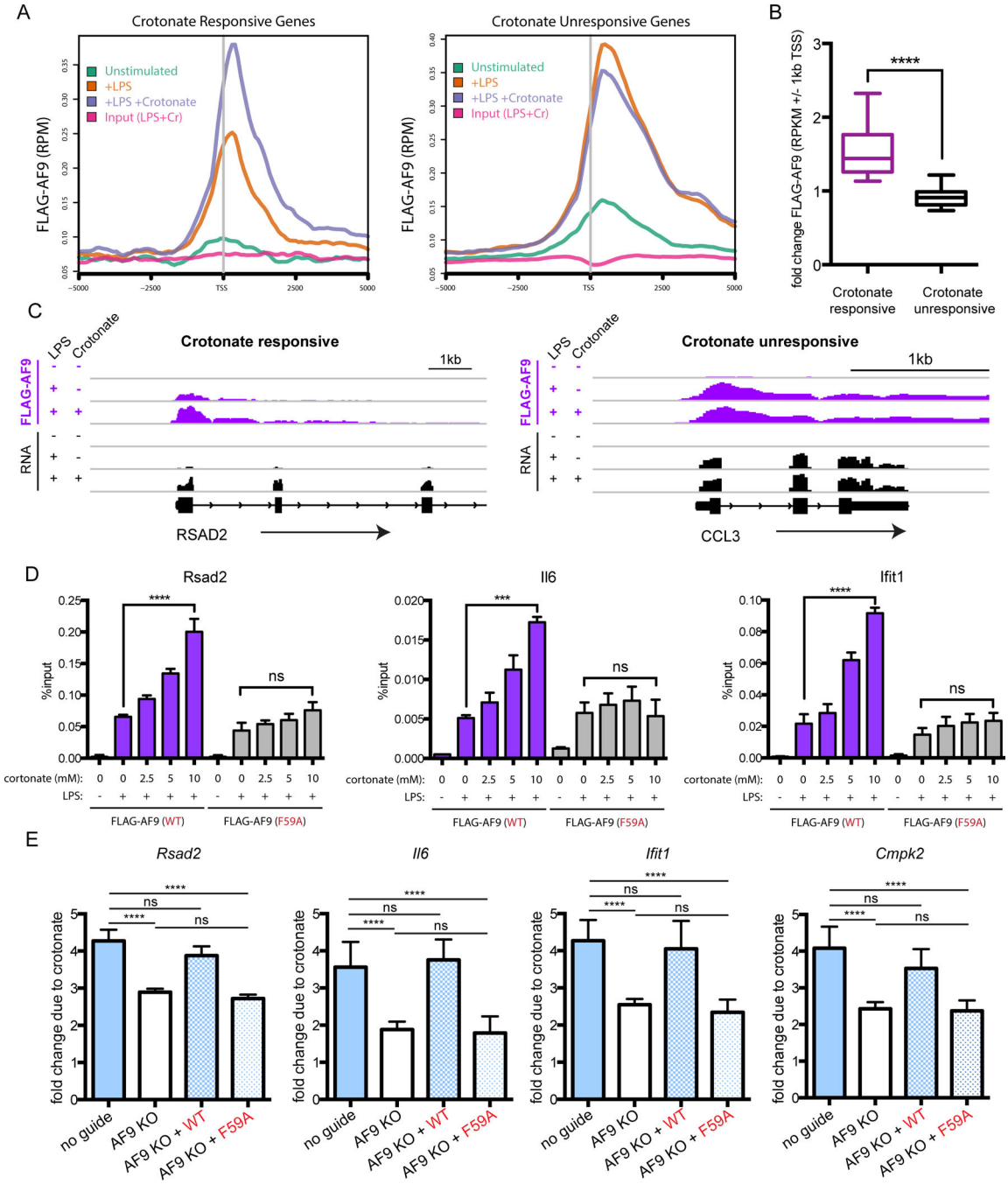


Figure 6. AF9 Co-Localizes with H3K18cr to Positively Regulate Gene Expression in a YEATS-Dependent Manner

(A) AF9 is further recruited to genes transcriptionally responsive to crotonate pre-treatment. Average profile of FLAG-AF9 ChIP-seq data from RAW264.7 cells unstimulated (green), LPS-stimulated (orange), or pre-treated with 10mM crotonate and then LPS-stimulated (purple) ± 5kb TSS of genes either responsive to crotonate ($\log_2$(fold-FPKM)>1.0; 52 genes) (left), or genes unresponsive to crotonate ($\log_2$(fold-FPKM)<0.01; 233 genes) (right).

(B) Box and whisker plot (10–90th percentile) of fold change due to crotonate pre-treatment in FLAG-AF9 RPKM ± 1kb TSS of crotonate responsive and crotonate unresponsive genes. ****: t-test derived p-value < 0.0001.

(C) Genome-browser view of FLAG-AF9 ChIP-seq (purple) and RNA-seq (black) data from cells unstimulated, LPS-stimulated, or crotonate pre-treated and LPS-stimulated for a representative responsive gene (*Rsad2*) and representative unresponsive gene (*Ccl3*).

(D) The increased recruitment of AF9 due to crotonate pre-treatment is YEATS-Kcr-dependent. FLAG-ChIP from RAW264.7 cells expressing a wild type AF9 (FLAG-AF9(WT)) construct or an AF9 construct with the YEATS-Kcr abrogating F59A mutation (FLAG-AF9(F59A)) under the conditions indicated followed by qPCR analysis of ChIP product and appropriate inputs. Data are plotted as mean percent input + standard deviation. ***: t-test derived p-value = 0.0001; ****: <0.0001; ns >0.05.

(E) The transcriptional response to increased crotonyl-CoA is dependent on AF9 and the YEATS-Kcr interaction. The fold changes in mRNA abundance for *Rsad2*, *Il6*, *Ifit1*, and *Cmpk2* due to crotonate pre-treatment prior to LPS stimulation are compared across four RAW264.7 cell lines: 1) wild type control (no guide), 2) CRISPR-Cas9 mediated AF9 genetic knockout (AF9 KO), 3) AF9 KO expressing a wild type construct of FLAG-AF9 (AF9KO + WT), and 4) AF9 KO expressing a F59A mutant construct of FLAG-AF9 (AF9KO + F59A). Data are plotted as mean fold change, as measured by qRT-PCR, of two biological replicates + standard deviation. ****: t-test derived p-value < 0.0001; ns: >0.05. See also Figure S6.

Table 1

Data collection and refinement statistics

Crystal	$AF9_{YEATS}$-H3K9cr	$AF9_{YEATS}$-H3K18cr	$BRD3_{BrD2}$-H3K18ac
Data Collection			
Space group	$P3_221$	$P2_1$	$P6_122$
Unit Cell			
a, *b*, *c* (Å)	97.6, 97.6, 39.7	91.3, 101.6, 102.1	79.8, 79.8, 95.3
α, β, γ(°)	90, 90, 120	90,103.1, 90	90, 90, 120
Resolution (Å)	50–2.7 (2.75–2.70)[a]	50–2.8 (2.85–2.80)	50–2.6 (2.64–2.60)
R_{sym} (%) [b]	6.6 (90.7)	11.6 (88.0)	18.6 (79.6)
I/Σ (I)	26.7 (2.8)	15.5 (2.7)	13.7 (3.8)
Completeness (%)	99.8 (99.0)	99.5 (95.5)	99.9 (100)
Redundancy	8.6 (8.6)	3.8 (3.8)	6.7 (7.1)
Wilson B-factor ($Å^2$)	74.1	52.6	33.2
Refinement (F>0)			
Resolution (Å)	35.9–2.7 (2.8–2.7)	48.3–2.8 (2.9–2.8)	28.6–2.6 (2.7–2.6)
No. of unique reflections	6,149	44,285	5,923
R_{work}/R_{free} (%) [c]	20.6/25.3	24.2/27.3	21.7/26.9
No. of non-H atoms			
Protein	1168	9216	913
Peptide	68	444	73
Water	16	41	52
Other Ligands	-	74	5
Average B-factors ($Å^2$)			
Protein	82.4	70.8	30.3
Peptide	81.0	85.6	43.4
Water	79.7	67.7	29.6
Other Ligands	-	94.5	41.2
RMSD bonds (Å)	0.010	0.004	0.002
RMSD angle (°)	1.27	0.84	0.64

[a] Numbers in parentheses refer to the highest resolution shells

[b] $R_{sym} = \Sigma_h\Sigma_i | I_i(h)-<I(h)> | / \Sigma_h\Sigma_i I_i(h)$, where $I_i(h)$ is the th measurement of reflection h, and <I(h)> is the weighted mean of all measurement of h.

[c] $R = \Sigma_h||F_{obs}|-|F_{cal}|| / \Sigma_h|F_{obs}|$, where F_{obs} and F_{cal} are the observed and calculated structure factor, respectively. R_{work} and R_{free} were calculated by using the working and test set reflections, respectively.