Published in final edited form as:

Biochim Biophys Acta. 2015 June ; 1849(6): 583–589. doi:10.1016/j.bbagr.2015.03.002.

C/EBP β (CEBPB) protein binding to the C/EBP|CRE DNA 8-mer TTGC|GTCA is inhibited by 5hmC and enhanced by 5mC, 5fC, and 5caC in the CG dinucleotide.

Syed Khund Sayeed¹, Jianfei Zhao¹, Bangalore K. Sathyanarayana², Jaya Prakash Golla¹, Charles Vinson^{1,*}

¹Laboratory of Metabolism, National Cancer Institute, National Institutes of Health, Room 3128, Building 37, Bethesda, MD 20892.

²Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Room 5120, Building 37, Bethesda, MD 20892.

Abstract

During mammalian development, some methylated cytosines (5mC) in CG dinucleotides are iteratively oxidized by TET dioxygenases to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). The effect of these cytosine oxidative products on the sequence-specific DNA binding of transcription factors is being actively investigated. Here, we used the electrophoretic mobility shift assay (EMSA) to examine C/EBP α and C/EBP β homodimers binding to all 25 chemical forms of a CG dinucleotide for two DNA sequences: the canonical C/EBP 8-mer TTGC|GCAA and the chimeric C/EBP|CRE 8-mer TTGC|GTCA. 5hmC in the CG dinucleotide in the C/EBP|CRE motif 8-mer TGAC|GCAA inhibits binding of C/EBP β but not C/EBP α . Binding was increased by 5mC, 5fC and 5caC. Circular dichroism monitored thermal denaturations for C/EBP β binding the C/EBP|CRE motif confirmed the EMSA. The structural differences between C/EBP α and C/EBP β that may account for the difference in binding 5hmC in the 8-mer TGAC|GCAA are explored.

Keywords

5mC; 5hmC; 5fC; 5caC; Carboxylation; C/EBP|CRE motif; CG dinucleotide; basic-leucine zipper; DNA binding; C/EBP β homodimer

1. Introduction

In mammals, most of the cytosines in CG dinucleotides are methylated [1]. Recently, the ten eleven translocation (TET) family of dioxygenases were identified [2,3] that iteratively oxidize 5mC into 5hmC [3] 5fC, and 5caC [4]. 5fC and 5caC can be removed by

*To whom correspondence should be addressed: Tel: 1-301-496-8783; Fax: 1-301-496-8419; vinsonc@mail.nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

mammalian thymine DNA glycosylase (TDG) and replaced by cytosine, completing the demethylation of 5mC [5,6] that occurs during many developmental stages [7], and physiological and pathological conditions [8], [9]. The abundance of the 5mC oxidative products varies in tissues [10] suggesting they are regulated intermediates with potential biological functions.

The effect of 5mC on the DNA binding of some transcription factors (TFs) has been examined. 5mC inhibits the DNA binding of many TFs involved in housekeeping functions like ETS (CCGGAA), SP1 (CCCGCC), and NRF-1 (CGCCTGCG) [11] suggesting a mechanistic link between hypermethylation of CG islands and gene suppression that is observed in some cancers [12,13]. Alternatively, 5mC can increase DNA binding of TFs [14] resulting in repression [15] or activation of nearby genes [16,17] [18] [19]. CG methylation improves binding of the C/EBP family of transcription factors [16,17,18] that are critical for activation of tissue specific promoters in many tissues during differentiation [20,21,22,23]. Furthermore, the C/EBP β |ATF4 heterodimer preferentially binds the methylated CGAT|GCAA [24].

A potential consequence of 5mC oxidation is to change the sequence-specific binding of transcription factors (TFs) [25,26,27,28]. In the present study, we used EMSA and CD thermal denaturations to examine the DNA binding of C/EBP α and C/EBP β to 25 double-stranded DNAs (dsDNA) containing all possible combinations of C, 5mC, 5hmC, 5fC, and 5caC in both cytosines of a CG dinucleotide for two DNA sequences, the canonical palindromic C/EBP 8-mer (TTGC|GCAA) [29,30] and the chimeric C/EBP|CRE 8-mer TTGC|GTCA [31,32,33]. C/EBP family members are widely expressed transcription factors that regulate cellular proliferation and differentiation with C/EBP α more involved in terminal differentiation [18,22,34,35,36,37]. Here we report that the binding of C/EBP β changes more than C/EBP α when 5mC is oxidized. The strongest change in binding is for C/EBP β and the chimeric C/EBP|CRE 8-mer suggesting a potential change in function as 5mC becomes oxidized during development [38], physiology [39], and pathology [40].

2. Material and methods

2.2 DNA oligonucleotides:

Twenty single-stranded DNA (ssDNA) cartridge-purified 28-mer oligonucleotides (both sense and anti-sense strands for C, 5mC, 5hmC, 5fC, and 5caC) were purchased from W. M. Keck Oligonucleotide Synthesis Facility, Yale to examine two DNA sequences: the C/EBP consensus motif (TTGC|GCAA) and chimeric C/EBP|CRE motif (TTGC|GTCA). These oligos were validated by capillary electrophoresis and gave single peaks for each oligo. Five 28-mer DNAs for the C/EBP consensus motif (CTGACCCATATTGC|GCAATCTGACTGAC) termed sense-strand (a) contained different versions of cytosine in bold (C, 5mC, 5hmC, 5fC, and 5caC). The C/EBP consensus motif is underlined and the center of the dyad is marked. Five 28-mer DNAs termed anti-sense strand (b) (GTCAGTCAGATTGC|GCAATATGGGTCAG) contained different versions of cytosine in bold. The chimeric C/EBP|CRE 28-mer on the sense-strand (a) is CTGACCCATATTGC|GTCATCTGACTGAC. The anti-sense strand was end-labeled with γ -³²P ATP (specific activity 5000 Ci/mmol, MP Biomedicals) using T4 polynucleotide kinase (New England

Biolabs), and was purified by the ProbeQuant G-50 micro column (GE Healthcare Biosciences). The double-stranded DNA (dsDNA) probes were generated by annealing the labeled anti-sense strand and unlabeled sense strand.

2.3 EMSA:

The DNA binding of C/EBP α and C/EBP β to 25 probes was analyzed by EMSA by incubating the protein homodimers with 32 P-labeled dsDNA (7 pM) in EMSA buffer (150 mM KCl, 12.5 mM K₂HPO₄-KH₂PO₄, pH 7.4, 1 mM DTT, 0.25 mM EDTA, 0.5 mg/mL BSA, 10% glycerol, 0.02 μ g/ μ L poly dIdC, 10 mM MgCl₂) in a final volume of the reaction 20 μ L. After 20 minutes of incubation at room temperature, protein DNA complex was separated by Polyacrylamide Gel Electrophoresis (PAGE) and autoradiographed.

2.4 Protein expression and purification:

The DNA binding B-ZIP domains of C/EBP α and C/EBP β were cloned into *E. coli* protein expression vectors and purified as described previously [41].

2.5 Circular Dichroism spectroscopy:

CD spectroscopy was performed using a Jasco J-720 spectropolarimeter and thermal denaturation curves were fitted [42].

2.6 Crystal structure of transcription factor C/EBP β homodimer:

The image of the X-ray structure of the C/EBP β homodimer bound to DNA (pdb2e42) was generated using the program Chimera <http://www.cgl.ucsf.edu/chimera/>.

3. Results

3.1 C/EBP β binding to 25 dsDNAs containing 5 different cytosines in the canonical C/EBP motif TTGC|GCAA:

5mC in the CG dinucleotide at the center of the consensus C/EBP motif (TTGC|GCAA) and the chimeric C/EBP|CRE (TTGC|GTCA) increases DNA binding of C/EBP α and C/EBP β proteins [16,26]. We extended this analysis and evaluated how the three oxidative products of 5mC affected C/EBP α and C/EBP β binding to these two DNA 8-mers with a CG dinucleotide at the center of the 8-mer. For each sequence, we designed 10 ssDNA 28-mers with the 8-mer motif in the center. Five ssDNAs have different cytosines (C, 5mC, 5hmC, 5fC and 5caC) in the CG dinucleotide in sense strand CTGACCCATATTGC|GCAATCTGACTGAC and 5 ssDNAs have different cytosines in the anti-sense strand GTCAGTCAGATTGC|GCAATATGGGTCAG. Sense and anti-sense ssDNAs were mixed to make 25 dsDNAs containing different chemical forms of the CG dinucleotide.

Figure 1A presents an EMSA with 3 nM C/EBP β homodimer binding to 25 radioactive dsDNAs containing different chemical forms of the CG dinucleotide in the palindromic consensus C/EBP motif 8-mer TTGC|GCAA. In dsDNA a CG dinucleotide contains two cytosines, thus the two cytosines may contain either the same or different modifications. For homotypic CG dinucleotide interactions, compared to C|C, 5mC|5mC, 5fC|5fC, and 5caC|5caC increase binding, whereas 5hmC|5hmC decreases binding. For heterotypic interactions,

5hmC|C and 5caC|C is better bound than C|C. The heterotypic CG dinucleotides appear as a sum of the contribution of each cytosine. We highlight the 10 pairs of DNA probes having heterotypic cytosines that are stereochemically equivalent within the palindromic canonical 8-mer and should have similar binding.

3.2 C/EBP β binding to 25 dsDNAs containing the chimeric C/EBP|CRE 8-mer TTGC|GTCA: GTCA:

Next we examined C/EBP β binding to the chimeric C/EBP|CRE motif TTGC|GTCA. Unlike the canonical palindromic C/EBP motif, the chimeric C/EBP|CRE 8-mer TTGC|GTCA is not palindromic and thus the two cytosines in the CG dinucleotide are stereochemically different and thus pairs of heterotypic cytosine interactions are not equivalent. C/EBP|CRE motif contains two B-ZIP half sites, each with a CG dinucleotide at the center of the dyad. The first half site (TTGC|C) is from the palindromic C/EBP motif (TTGC|GCAA) discussed above. The second half site (C|GTCA) is from the palindromic CRE (TGAC|GTCA) bound by the CREB homodimer and other B-ZIP proteins [43].

Figure 1B presents an EMSA with 10 nM C/EBP β homodimer binding to all 25 chemical forms of the CG dinucleotide in the C/EBP|CRE chimeric 8-mer TTGC|GTCA. A three-fold increase in protein concentration was needed because of weaker C/EBP β binding to the chimeric site compared to the canonical C/EBP motif. Cytosine modifications change C/EBP β binding to the chimeric motif (TTGC|GTCA) more than the C/EBP motif (TTGC|GCAA). 5mC modestly increased binding, 5hmC decreased binding, and both 5fC and 5caC increased binding of C/EBP β binding to the C/EBP|CRE motif.

Figure 1C presents an EMSA for C/EBP β homodimer binding using a half-log dilution from 100 nM to 3 nM to three dsDNAs: unmodified DNA, dsDNA with two 5mCs, and dsDNA with two 5hmCs in the CG dinucleotide in the canonical C/EBP motif TTGC|GCAA. 5mC increases binding of C/EBP β homodimer while 5hmC decreases binding compared to unmodified cytosine. Figure 1D presents an EMSA of the C/EBP β homodimer, using a half-log dilution from 100 nM to 3 nM, binding to four dsDNAs containing the chimeric C/EBP|CRE motif showing that methylation of the CG dinucleotide increases binding and that oxidation of either cytosine to 5hmC inhibits binding compared to unmodified cytosine.

3.3 C/EBP α binding to 25 dsDNA for two motifs:

3 nM C/EBP α homodimer binding to all 25 CG chemical forms of C/EBP motif 8-mer TTGC|GCAA (Figure 2A) is similar to C/EBP β (Figure 1A). However, 10 nM C/EBP α and C/EBP β homodimer binding to all 25 CG chemical forms of the C/EBP|CRE chimeric 8-mer TTGC|GTCA is different with 5hmC inhibiting the binding of C/EBP β but not C/EBP α (Figure 2B).

Figure 2C presents an EMSA for C/EBP α homodimer using a half-log dilution from 100 nM to 3 nM binding to two DNAs, an unmodified DNA and one with two 5caCs in the CG dinucleotide in TTGC|GCAA and shows that 5caC increases binding of C/EBP α homodimer by ~ 4-fold.

3.4 Circular dichroism (CD) monitored thermal stability of C/EBP β bound to chimeric C/EBP|CRE motif (TTGC|GTCA):

We used thermal stability monitored by CD spectroscopy as an independent method to evaluate C/EBP β binding to modified C/EBP|CRE motifs. Initially, we determined the thermal stability of 16 dsDNAs 28-mers containing different combination of cytosine and modified cytosine using 245 nm. Similar results to our previous study using an E-Box motif were obtained [25]. Among the homotypic dsDNA, 5mC|5mC, 5hmC|5hmC, and 5fC|5fC modestly increase stability whereas 5caC|5caC decreases stability. dsDNAs with only one 5caC has intermediate stability between C|C and 5caC|5caC suggesting the modifications of the two cytosines in the CG dinucleotide are acting independently (Table 1).

Next we examined C/EBP β homodimer thermal denaturation monitored by CD spectroscopy at 222 nm as an independent method to assess binding to the chimeric C/EBP|CRE DNA motif. Figure 3A presents CD spectra at 6° C from 200 nm to 300 nm of three samples 1) dsDNA 28-mer containing the chimeric C/EBP|CRE motif TTGC|GTCA, 2) the C/EBP β homodimer, and 3) the C/EBP β homodimer bound to dsDNA 28-mer. An increase in negative ellipticity is observed at 222 nm with the addition of DNA as has been observed previously [41] and interpreted as an increase in α -helical structure as the basic region become helical in the major groove of DNA [44,45].

Figure 3B presents the thermal stability of dsDNA monitored at 245 nm with a T_m of 69.6° C. The C/EBP β homodimer monitored at 222 nm has a T_m of 61.6° C. With heating, the C/EBP β homodimer cooperatively losses ellipticity at 222 nm as observed previously [41]. The denaturation is well fit using a two-state model of α -helical dimers becoming un-helical monomers. The addition of dsDNA did not dramatically change the thermal stability of the C/EBP β homodimer (T_m 60.8° C), unlike other B-ZIP dimers where the addition of DNA increases protein stability [16].

We next examined C/EBP β homodimer binding to 16 dsDNA. Among the homotypic dsDNA, compared to unmodified cytosine 5mC, 5fC and 5caC modestly increase stability of the protein-DNA complex whereas 5hmC decreases the stability, recapitulating what was observed in EMSA (Table1).

3.4 Molecular modeling of C/EBP β bound to 5mC dsDNA.

To evaluate the structural underpinnings for the differential binding between C/EBP α and C/EBP β to 5hmC containing C/EBP|CRE motif, we examined an 1.8 angstrom crystal structure of C/EBP β bound to a canonical DNA motif (pdb 2e42) and highlight the modified cytosine and the amino acids that are different between C/EBP α and C/EBP β . The closest is Met 294 which is 14 angstroms from 5mC suggesting that the differences in binding are subtle. We presume that the differences in binding affinities of C/EBP α and C/EBP β are due to the differences in the amino-acids near the modified cytosine.

4. Discussion

The potential for 5mC, 5hmC, 5fC, and 5caC to change the sequence specific DNA binding of transcription factors adds complexity to understanding regulated gene expression [6]. The

effect of 5mC on sequence-specific DNA binding has been examined for many transcription factors [24] [14] but how the oxidative products of 5mC change DNA binding is less understood. Here, we show that these modifications have modest effects on C/EBP α binding to the palindromic consensus motif (TTGC|GCAA) and the chimeric C/EBP|CRE motif (TTGC|GTCA). In contrast, the DNA binding of the C/EBP β homodimer is more sensitive to these changes with 5mC [16,24], 5fC, and 5caC increasing binding while 5hmC decreases binding. Introduction of 5hmC results in widening of the major groove at the site of modification[28] inhibit B-Z transition of DNA [46]. 5hmC has been shown to be repressive for transcription [47,48,49]. Interestingly, presence of 5hmC completely abrogates the binding of BHLH protein MAX, USF and HIF1[28]. Alteration of DNA structures has also been reported for 5fC [50] and 5caC[46]. A proteomic screen for proteins binding cytosine and its different modifications has identified novel proteins that bind to 5fC and 5caC with potential regulatory functions [51].

5fC preferentially accumulates at poised enhancers among other gene regulatory elements in mouse embryonic stem cells [52,53,54]. Potentially, C/EBP β preferential binding to 5fC and 5caC containing C/EBP|CRE motif is important for the regulatory events that occur as parts of the genome become unmethylated. The decrease in C/EBP β binding to 5hmC containing C/EBP|CRE motif may be significant as this modification accumulates in tissues, particularly the brain [10] [40] [55]. It is plausible that 5hmC may act as a switch by inhibiting C/EBP β binding and thus promoting C/EBP α binding during differentiation of brain tissue. Transient expression of C/EBP α and C/EBP β induce Tet2 expression in mouse B cells and facilitate its nuclear translocation upon induction [56] suggesting that C/EBP family members recruit Tet dioxygenases to 5mC containing motifs resulting in oxidation and changes in C/EBP family binding which is critical for the complex changes that occur with demethylation and gene expression.

The methods exist to answer if 5hmC changes the occupancy of C/EBP α and C/EBP β binding to the modified C/EBP|CRE motif. By performing ChIP with C/EBP α and C/EBP β antibodies followed by the determination of the methylation and hydroxymethylation status [57], one can answer if C/EBP α is enriched in 5hmC containing DNA. Additionally one can determine the enrichment of 5fC and 5caC in C/EBP α and C/EBP β ChIP DNA [54]. Our work supports the notion/hypothesis that cytosine modifications regulate transcription in a sequence and context dependent manner and lays the foundation for further exploring the complex interactions between C/EBP family members and 5mC and its oxidative products.

Acknowledgments:

We thank our lab members for their encouragement and support. This work is supported by the intramural research project of National Cancer Institute, NIH, Bethesda, USA.

Abbreviations

C	Cytosine
B-ZIP	basic-leucine zipper
5mC	5-methylcytosine

5hmC	5-hydroxymethylcytosine
5fC	5-formylcytosine
5caC	5-carboxylcytosine
EMSA	electrophoretic mobility shift assay
dsDNA	double-stranded DNA
ssDNA	single-stranded DNA

References

- [1]. Deaton AM, Bird A, CpG islands and the regulation of transcription, *Genes Dev* 25 (2011) 1010–1022. [PubMed: 21576262]
- [2]. Iyer LM, Tahiliani M, Rao A, Aravind L, Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids, *Cell Cycle* 8 (2009) 1698–1710. [PubMed: 19411852]
- [3]. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A, Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1, *Science* 324 (2009) 930–935. [PubMed: 19372391]
- [4]. Cadet J, Wagner JR, TET enzymatic oxidation of 5-methylcytosine, 5-hydroxymethylcytosine and 5-formylcytosine, *Mutat Res* (2013).
- [5]. Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S, Cheng AW, Gao Q, Kim J, Choi SW, Page DC, Jaenisch R, Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development, *Cell Stem Cell* 9 (2011) 166–175. [PubMed: 21816367]
- [6]. Pastor WA, Aravind L, Rao A, TETonic shift: biological roles of TET proteins in DNA demethylation and transcription, *Nat Rev Mol Cell Biol* 14 (2013) 341–356. [PubMed: 23698584]
- [7]. Zhu S, Yoon K, Sterneck E, Johnson PF, Smart RC, CCAAT/enhancer binding protein-beta is a mediator of keratinocyte survival and skin tumorigenesis involving oncogenic Ras signaling, *Proc Natl Acad Sci U S A* 99 (2002) 207–212. [PubMed: 11756662]
- [8]. Ramji DP, Foka P, CCAAT/enhancer-binding proteins: structure, function and regulation, *Biochem J* 365 (2002) 561–575. [PubMed: 12006103]
- [9]. Liu Y, Nonnemacher MR, Wigdahl B, CCAAT/enhancer-binding proteins and the pathogenesis of retrovirus infection, *Future Microbiol* 4 (2009) 299–321. [PubMed: 19327116]
- [10]. Kriaucionis S, Heintz N, The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain, *Science* 324 (2009) 929–930. [PubMed: 19372393]
- [11]. Rozenberg JM, Shlyakhtenko A, Glass K, Rishi V, Myakishev MV, FitzGerald PC, Vinson C, All and only CpG containing sequences are enriched in promoters abundantly bound by RNA polymerase II in multiple tissues, *BMC Genomics* 9 (2008) 67. [PubMed: 18252004]
- [12]. Iguchi-Ariga SM, Schaffner W, CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation, *Genes Dev* 3 (1989) 612–619. [PubMed: 2545524]
- [13]. Tate PH, Bird AP, Effects of DNA methylation on DNA-binding proteins and gene expression, *Curr Opin Genet Dev* 3 (1993) 226–231. [PubMed: 8504247]
- [14]. Hu S, Wan J, Su Y, Song Q, Zeng Y, Nguyen HN, Shin J, Cox E, Rho HS, Woodard C, Xia S, Liu S, Lyu H, Ming GL, Wade H, Song H, Qian J, Zhu H, DNA methylation presents distinct binding sites for human transcription factors, *Elife* 2 (2013) e00726. [PubMed: 24015356]
- [15]. Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, Baglivo I, Pedone PV, Grimaldi G, Riccio A, Trono D, In embryonic stem cells, ZFP57/KAP1 recognize a

methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions, *Mol Cell* 44 (2011) 361–372. [PubMed: 22055183]

- [16]. Rishi V, Bhattacharya P, Chatterjee R, Rozenberg J, Zhao J, Glass K, Fitzgerald P, Vinson C, CpG methylation of half-CRE sequences creates C/EBPalpha binding sites that activate some tissue-specific genes, *Proc Natl Acad Sci U S A* 107 (2010) 20311–20316. [PubMed: 21059933]
- [17]. Flower K, Thomas D, Heather J, Ramasubramanyan S, Jones S, Sinclair AJ, Epigenetic control of viral life-cycle by a DNA-methylation dependent transcription factor, *PLoS One* 6 (2011) e25922. [PubMed: 22022468]
- [18]. Vinson C, Chatterjee R, CG methylation, *Epigenomics* 4 (2012) 655–663. [PubMed: 23244310]
- [19]. Gustems M, Woellmer A, Rothbauer U, Eck SH, Wieland T, Lutter D, Hammerschmidt W, c-Jun/c-Fos heterodimers regulate cellular genes via a newly identified class of methylated DNA sequence motifs, *Nucleic Acids Res* 42 (2014) 3059–3072. [PubMed: 24371273]
- [20]. Buck M, Poli V, Hunter T, Chojkier M, C/EBPbeta phosphorylation by RSK creates a functional XEXD caspase inhibitory box critical for cell survival, *Mol Cell* 8 (2001) 807–816. [PubMed: 11684016]
- [21]. Darlington GJ, Ross SE, MacDougald OA, The role of C/EBP genes in adipocyte differentiation, *J Biol Chem* 273 (1998) 30057–30060. [PubMed: 9804754]
- [22]. Smink JJ, Leutz A, Rapamycin and the transcription factor C/EBPbeta as a switch in osteoclast differentiation: implications for lytic bone diseases, *J Mol Med (Berl)* 88 (2010) 227–233. [PubMed: 19943027]
- [23]. Tanaka T, Yoshida N, Kishimoto T, Akira S, Defective adipocyte differentiation in mice lacking the C/EBPbeta and/or C/EBPdelta gene, *EMBO J* 16 (1997) 7432–7443. [PubMed: 9405372]
- [24]. Mann IK, Chatterjee R, Zhao J, He X, Weirauch MT, Hughes TR, Vinson C, CG methylated microarrays identify a novel methylated sequence bound by the CEBPB|ATF4 heterodimer that is active in vivo, *Genome Res* 23 (2013) 988–997. [PubMed: 23590861]
- [25]. Golla JP, Zhao J, Mann IK, Sayeed SK, Mandal A, Rose RB, Vinson C, Carboxylation of cytosine (5caC) in the CG dinucleotide in the E-box motif (CGCAG|GTG) increases binding of the Tcf3|Ascl1 helix-loop-helix heterodimer 10-fold, *Biochem Biophys Res Commun* (2014).
- [26]. Chatterjee R, Vinson C, CpG methylation recruits sequence specific transcription factors essential for tissue specific gene expression, *Biochim Biophys Acta* 1819 (2012) 763–770. [PubMed: 22387149]
- [27]. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PW, Bauer C, Munzel M, Wagner M, Muller M, Khan F, Eberl HC, Mensinga A, Brinkman AB, Lephikov K, Muller U, Walter J, Boelens R, van Ingen H, Leonhardt H, Carell T, Vermeulen M, Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives, *Cell* 152 (2013) 1146–1159. [PubMed: 23434322]
- [28]. Lercher L, McDonough MA, El-Sagheer AH, Thalhammer A, Kriaucionis S, Brown T, Schofield CJ, Structural insights into how 5-hydroxymethylation influences transcription factor binding, *Chem Commun (Camb)* 50 (2014) 1794–1796. [PubMed: 24287551]
- [29]. Johnson PF, Identification of C/EBP basic region residues involved in DNA sequence recognition and half-site spacing preference, *Mol Cell Biol* 13 (1993) 6919–6930. [PubMed: 8413284]
- [30]. D. I, An integrated encyclopedia of DNA elements in the human genome, *Nature* 489 (2012) 57–74. [PubMed: 22955616]
- [31]. Vinson CR, Hai T, Boyd SM, Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: prediction and rational design, *Genes Dev* 7 (1993) 1047–1058. [PubMed: 8504929]
- [32]. Vallejo M, Ron D, Miller CP, Habener JF, C/ATF, a member of the activating transcription factor family of DNA-binding proteins, dimerizes with CAAT/enhancer-binding proteins and directs their binding to cAMP response elements, *Proc Natl Acad Sci U S A* 90 (1993) 4679–4683. [PubMed: 8506317]
- [33]. FitzGerald PC, Shlyakhtenko A, Mir AA, Vinson C, Clustering of DNA sequences in human promoters, *Genome Res* 14 (2004) 1562–1574. [PubMed: 15256515]
- [34]. Hata K, Nishimura R, Ueda M, Ikeda F, Matsubara T, Ichida F, Hisada K, Nokubi T, Yamaguchi A, Yoneda T, A CCAAT/enhancer binding protein beta isoform, liver-enriched inhibitory protein,

regulates commitment of osteoblasts and adipocytes, *Mol Cell Biol* 25 (2005) 1971–1979. [PubMed: 15713650]

- [35]. Nerlov C, The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control, *Trends Cell Biol* 17 (2007) 318–324. [PubMed: 17658261]
- [36]. Hirata M, Kugimiya F, Fukai A, Ohba S, Kawamura N, Ogasawara T, Kawasaki Y, Saito T, Yano F, Ikeda T, Nakamura K, Chung UI, Kawaguchi H, C/EBPbeta Promotes transition from proliferation to hypertrophic differentiation of chondrocytes through transactivation of p57, *PLoS One* 4 (2009) e4543. [PubMed: 19229324]
- [37]. Stoilova B, Kowenz-Leutz E, Scheller M, Leutz A, Lymphoid to myeloid cell trans-differentiation is determined by C/EBPbeta structure and post-translational modifications, *PLoS One* 8 (2013) e65169. [PubMed: 23755188]
- [38]. Tsagaratou A, Aijo T, Lio CW, Yue X, Huang Y, Jacobsen SE, Lahdesmaki H, Rao A, Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation, *Proc Natl Acad Sci U S A* (2014).
- [39]. Hackett JA, Dietmann S, Murakami K, Down TA, Leitch HG, Surani MA, Synergistic Mechanisms of DNA Demethylation during Transition to Ground-State Pluripotency, *Stem Cell Reports* 1 (2013) 518–531. [PubMed: 24371807]
- [40]. Pfeifer GP, Xiong W, Hahn MA, Jin SG, The role of 5-hydroxymethylcytosine in human cancer, *Cell Tissue Res* 356 (2014) 631–641. [PubMed: 24816989]
- [41]. Rishi V, Potter T, Laudeman J, Reinhart R, Silvers T, Selby M, Stevenson T, Krosky P, Stephen AG, Acharya A, Moll J, Oh WJ, Scudiero D, Shoemaker RH, Vinson C, A high-throughput fluorescence-anisotropy screen that identifies small molecule inhibitors of the DNA binding of B-ZIP transcription factors, *Anal Biochem* 340 (2005) 259–271. [PubMed: 15840499]
- [42]. Krylov D, Mikhailenko I, Vinson C, A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions, *EMBO J* 13 (1994) 2849–2861. [PubMed: 8026470]
- [43]. Vinson C, Myakishev M, Acharya A, Mir AA, Moll JR, Bonovich M, Classification of Human B-ZIP Proteins Based on Dimerization Properties, *Mol Cell Biol* 22 (2002) 6321–6335. [PubMed: 12192032]
- [44]. Vinson CR, LaMarco KL, Johnson PF, Landschulz WH, McKnight SL, In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage, *Genes Dev* 2 (1988) 801–806. [PubMed: 3061875]
- [45]. Shuman JD, Vinson CR, McKnight SL, Evidence of changes in protease sensitivity and subunit exchange rate on DNA binding by C/EBP, *Science* 249 (1990) 771–774. [PubMed: 2202050]
- [46]. Wang S, Long Y, Wang J, Ge Y, Guo P, Liu Y, Tian T, Zhou X, Systematic Investigations of Different Cytosine Modifications on CpG Dinucleotide Sequences: The Effects on the B-Z Transition, *Journal of the American Chemical Society* 136 (2013) 56–59. [PubMed: 24364741]
- [47]. Robertson J, Robertson AB, Klungland A, The presence of 5-hydroxymethylcytosine at the gene promoter and not in the gene body negatively regulates gene expression, *Biochem Biophys Res Commun* 411 (2011) 40–43. [PubMed: 21703242]
- [48]. Choi I, Kim R, Lim HW, Kaestner KH, Won KJ, 5-hydroxymethylcytosine represses the activity of enhancers in embryonic stem cells: a new epigenetic signature for gene regulation, *BMC Genomics* 15 (2014) 670. [PubMed: 25106691]
- [49]. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J, Helin K, TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity, *Nature* 473 (2011) 343–348. [PubMed: 21490601]
- [50]. Raiber EA, Murat P, Chirgadze DY, Beraldi D, Luisi BF, Balasubramanian S, 5-Formylcytosine alters the structure of the DNA double helix, *Nat Struct Mol Biol* 22 (2015) 44–49. [PubMed: 25504322]
- [51]. Iurlaro M, Ficiz G, Oxley D, Raiber EA, Bachman M, Booth MJ, Andrews S, Balasubramanian S, Reik W, A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation, *Genome Biol* 14 (2013) R119. [PubMed: 24156278]

- [52]. Song CX, Szulwach KE, Dai Q, Fu Y, Mao SQ, Lin L, Street C, Li Y, Poidevin M, Wu H, Gao J, Liu P, Li L, Xu GL, Jin P, He C, Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming, *Cell* 153 (2013) 678–691. [PubMed: 23602153]
- [53]. Cadet J, Wagner JR, TET enzymatic oxidation of 5-methylcytosine, 5-hydroxymethylcytosine and 5-formylcytosine, *Mutat Res Genet Toxicol Environ Mutagen* 764–765 (2014) 18–35.
- [54]. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, Zhang K, Zhang Y, Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics, *Cell* 153 (2013) 692–706. [PubMed: 23602152]
- [55]. Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, Xie ZG, Shi L, He X, Jin SG, Iqbal K, Shi YG, Deng Z, Szabo PE, Pfeifer GP, Li J, Xu GL, The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes, *Nature* 477 (2011) 606–610. [PubMed: 21892189]
- [56]. Di Stefano B, Sardina JL, van Oevelen C, Collombet S, Kallin EM, Vicent GP, Lu J, Thieffry D, Beato M, Graf T, C/EBPalpha poises B cells for rapid reprogramming into induced pluripotent stem cells, *Nature* 506 (2014) 235–239. [PubMed: 24336202]
- [57]. Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A, Li X, Dai Q, Shen Y, Park B, Min JH, Jin P, Ren B, He C, Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome, *Cell* 149 (2012) 1368–1380. [PubMed: 22608086]

Highlights of paper

1. We examined effect of 5mC, 5hmC, 5fC and 5caC on DNA binding of C/EBP α and C/EBP β .
2. DNA-binding of C/EBP β changes more than C/EBP α when 5mC is oxidized in TGAC|GCAA.
3. 5mC, 5fC and 5caC in CG dinucleotide of TGAC|GCAA enhance DNA binding of C/EBP β .
4. 5hmC in CG dinucleotide of TGAC|GCAA inhibits binding of C/EBP β but not C/EBP α .

Figure 1: C/EBPβ homodimer binding DNA containing modified cytosines:

A) Binding of 3 nM C/EBPβ homodimer to 25 dsDNA containing different cytosine forms in the CG dinucleotide in the canonical C/EBP motif TTGC|GCAA. The five homotypic interactions are indicated by * and the ten heterotypic interactions are numbered 1 to 10. **B)** Binding of 10 nM C/EBPβ homodimer to 25 dsDNA containing different cytosine forms in the CG dinucleotide in the chimeric C/EBP|CRE motif TTGC|GTCA. **C)** EMSA showing a half-log dilution from 100 nM to 3 nM for the C/EBPβ homodimer binding three dsDNA containing the C/EBP motif (TTGC|GCAA), one contains unmodified cytosine, one contains two 5mC, and the last contain two 5hmC. **D)** EMSA showing a half-log dilution from 100 nM to 3 nM for the C/EBPβ homodimer binding four dsDNA containing the chimeric C/EBP|CRE motif (TTGC|GTCA) at its center. One contains unmodified cytosine, one contains two 5mCs, and the last two dsDNA contain both possible 5hmC and 5mC combinations in the CG dinucleotide

Figure 2: C/EBPα homodimer binding DNA containing modified cytosines:

A) Binding of 3 nM C/EBPα homodimer to 25 dsDNA containing different cytosine forms in the CG dinucleotide in the canonical C/EBP motif TTGC|GCAA. The five homotypic interactions are indicated by * and the ten heterotypic interactions are numbered 1 to 10. **B)** Binding of 10 nM C/EBPα homodimer to 25 dsDNA containing different cytosine forms in the CG dinucleotide in the chimeric C/EBP|CRE motif TTGC|GTCA. **C)** EMSA showing a half-log dilution from 100 nM to 3 nM for the C/EBPα homodimer binding to two dsDNA containing the canonical C/EBP motif (TTGC|GCAA), one dsDNA contains unmodified cytosine, and the other dsDNA contains two 5caCs.

Figure 3:

A) CD spectra from 200 nm to 300 nm at 6°C for C/EBPβ, C/EBPβ homodimer bound to a dsDNA 28-mer and unmodified dsDNA 28-mer containing the C/EBP|CRE chimeric motif TTGC|GTCA. **B)** Circular dichroism at 222 nm of the thermal denaturation at 2 μM dimer of three samples: C/EBPβ (*), C/EBPβ homodimer bound to a dsDNA 28-mer containing the C/EBP|CRE chimeric motif TTGC|GTCA (□). Thermal denaturation at 245 nm of 2 μM of dsDNA 28-mer containing the C/EBP|CRE chimeric motif TTGC|GTCA (○).

(A)

	Basic region	Leucine zipper
C/EBP α	SVDK NSNEYRV RRERNNIAVRKSRDKAK Q RNVET Q	KVLELTSDNDRLRK–
C/EBP β	T VDK HSDEYKI RRERNNIAVRKSRDKAK MRN LET QH	KVLELTAENERLQK–

(B)

Figure 4.

A) Sequence alignment of C/EBP α and C/EBP β showing the differences in the amino-acids between C/EBP α and C/EBP β . B). Molecular model of C/EBP β homodimer bound to consensus DNA sequence. The residues shown in colored spheres represent amino-acids that are different from C/EBP α . 5hmC in one of the DNA strands is shown in magenta.

Table1:
Thermal stability of DNA and C/EBP β bound to DNA containing cytosine modifications:

Of the 25 forms of the CG dinucleotide, we analyzed 17 forms. 5 had homotypic interactions (C|C, 5mC|5mC, 5hmC|5hmC, 5fC|5fC, and 5caC|5caC; represented by * in the EMSAs) and the remaining 12 had heterotypic interactions.

	28-mer dsDNA stability (245 nm) TGAC CGAA(b)					
	T _m (°C) ± SE	C(b)	5mC(b)	5hmC(b)	5fC(b)	5caC(b)
C/EBP CRE: TTGC GTCA (a)	C(a)	69.6±0.39	69.3±0.38	69.9±0.33		67.1±0.35
	5mC(a)	68.4±0.40	70.6±0.27	69.9±0.22	70.3±0.31	68.6±0.39
	5hmC(a)	68.9±0.27	70.3±0.22	71.1±0.16		
	5fC(a)		70.7±0.21		70.6±0.19	
	5caC(a)	68.7±0.37				66.9±0.3
	C/EBP β stability bound to dsDNA (222 nm) TGAC CGAA(b)					
	C(a)	60.8±0.11	62.1±0.09	60.5±0.09		62.8±0.09
	5mC(a)	61.5±0.11	61.5±0.09	60.8±0.10	62.5±0.09	63.5±0.08
	5hmC(a)	59.1±0.15	60.7±0.12	58.5±0.15		
	5fC(a)		63.0±0.08		62.4±0.09	
	5caC(a)	62.2±0.10				61.9±0.11