



<https://doi.org/10.1038/s42003-022-03099-0>

OPEN

H3K27ac nucleosomes facilitate HMGN localization at regulatory sites to modulate chromatin binding of transcription factors

Shaofei Zhang^{1,5}, Yuri Postnikov^{1,5}, Alexei Lobanov^{2,3}, Takashi Furusawa¹, Tao Deng^{1,4} & Michael Bustin¹  

Nucleosomes containing acetylated H3K27 are a major epigenetic mark of active chromatin and identify cell-type specific chromatin regulatory regions which serve as binding sites for transcription factors. Here we show that the ubiquitous nucleosome binding proteins HMGN1 and HMGN2 bind preferentially to H3K27ac nucleosomes at cell-type specific chromatin regulatory regions. HMGNs bind directly to the acetylated nucleosome; the H3K27ac residue and linker DNA facilitate the preferential binding of HMGNs to the modified nucleosomes. Loss of HMGNs increases the levels of H3K27me3 and the histone H1 occupancy at enhancers and promoters and alters the interaction of transcription factors with chromatin. These experiments indicate that the H3K27ac epigenetic mark enhances the interaction of architectural protein with chromatin regulatory sites and identify determinants that facilitate the localization of HMGN proteins at regulatory sites to modulate cell-type specific gene expression.

¹Protein Section, Laboratory of Metabolism, Center for Cancer Research, National Cancer Institute, Maryland, USA. ²CCR Collaborative Bioinformatics Resource, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. ³Advanced Biomedical Computational Science, Frederick National Laboratory for Cancer Research, Maryland, MD, USA. ⁴Cell Translation Laboratory, NCATS, National Institutes of Health, 9800 Medical Center Drive, Rockville, MD 20850, USA. ⁵These authors equally contributed: Shaofei Zhang, Yuri Postnikov. ✉email: bustinm@mail.nih.gov

Correct binding of transcription factors (TFs) to their specific DNA motifs in chromatin plays a key role in establishing an epigenetic landscape that facilitates cell-type-specific gene expression necessary for the maintenance of cell identity^{1,2}. The interaction of TFs with chromatin is dynamic; TFs continuously move throughout the nucleus and reside temporarily at their specific binding sites^{3–5}. The binding of TFs to chromatin is facilitated by nuclear factors that help TFs access their binding sites and perturbed by factors that impede the binding of TFs to specific chromatin sites, especially at enhancers and promoters⁶, chromatin regions enriched in H3K27ac nucleosomes⁷. Possible regulators of TF chromatin binding include architectural proteins, such as the linker histone H1 and high mobility group (HMG) proteins; these ubiquitous structural proteins are known to affect chromatin organization in many cell types. Histone H1 facilitates the formation of higher-order chromatin organization and stabilizes compact chromatin structures that inhibit transcription^{8–10}, while HMG proteins are mostly associated with reduced chromatin compaction and increased gene expression from chromatin templates^{11,12}. Given the global effects of H1 and HMGs on chromatin structure and gene expression, it is likely that these ubiquitous structural proteins do modulate the binding of TFs to chromatin, a possibility that has not been studied in detail. Here we focus on the high mobility group N (HMGN) proteins and show that the major members of this family, HMGN1 and HMGN2, bind preferentially to nucleosomes containing the H3K27ac epigenetic mark, and affect the binding of TFs to chromatin.

HMGN is a family of abundant and evolutionarily conserved proteins that bind to nucleosomes without specificity for the underlying DNA sequence^{13,14}. The interaction of HMGN proteins with chromatin is highly dynamic; HMGNs bind to nucleosomes with short residence times and can be readily dislocated from their chromatin binding sites¹⁵. The binding of HMGNs to nucleosomes reduces chromatin compaction, most likely because it alters the interactions of linker histone H1 with chromatin^{16–18} and binds to the nucleosome acidic patch, a region thought to stabilize interactions between neighboring nucleosomes^{19,20}. Although HMGNs bind to chromatin without DNA sequence specificity, they preferentially localize to enhancers and promoters, chromatin regulatory sites that are easily digested with DNaseI and enriched in epigenetic marks of active chromatin, including H3K27ac modified histones^{21,22}. Changes in HMGN levels are frequently associated with a wide range of cell-type-specific changes in gene expression and phenotypes²³. Genetically altered mice lacking both HMGN1 and HMGN2 proteins (DKO mice) are born and survive but show multiple phenotypes²¹, (<https://www.mouseclinic.de/>). MEFs isolated from DKO mice can be reprogrammed into pluripotent cells by exogenous TFs more efficiently than MEFs isolated from WT mice suggesting that loss of HMGNs destabilized the maintenance of cell identity^{22,24}. In Down syndrome, one of the most prevalent human genetic diseases, the presence of an extra copy of *HMGN1* has been directly linked to increased levels of H3K27ac and to gene expression changes and to increased incidence of acute lymphoblastic leukemia^{25,26}. Taken together, the available data suggest that HMGNs modulate and fine-tune cell-type-specific gene expression programs.

Because the amount of HMGN in a cell is enough to bind only about 1% of the nucleosomes²⁷, it is not clear how these structural proteins, that interact with chromatin without DNA-sequence specificity, can nevertheless affect gene expression and cellular phenotypes in a cell-type-specific manner. Likely, HMGNs affect cell-type-specific gene expression by preferentially localizing to chromatin regulatory regions, as indicated by genome-wide analyses that show high HMGN occupancy at chromatin regions containing high levels of H3K27ac^{21,22}, an epigenetic modification

that marks promoters and enhancers^{7,28,29}. Conceivably, the presence of HMGN at enhancers and promoters may affect the interaction of TFs with these sites, thereby impacting cell-type-specific gene expression. The factors that target HMGN to enhancers and promoters and the possible effect of these proteins on TFs chromatin binding have not yet been investigated.

Here we identify the major determinants that facilitate the preferential binding of HMGN proteins to chromatin regulatory sites and show that HMGNs affect the binding of transcription factors to chromatin. Using bioinformatic analyses we demonstrate that HMGN1 and HMGN2 preferentially localize to nucleosomes containing the H3K27ac residue. We show that the presence of H3K27ac, an epigenetic mark of active chromatin, strengthens the binding of both HMGN1 and HMGN2 to the modified nucleosomes, and that loss of HMGNs alters H3K27 modifications and H1 occupancy at enhancers and promoters. We analyze the genome-wide binding of several TFs in cells derived from WT and DKO mice and find that loss of HMGNs alters the binding of TFs to chromatin, especially at enhancers. Our studies provide insights into factors that affect the recruitment of ubiquitous architectural chromatin binding proteins to regulatory chromatin to modulate transcription factor accessibility and fine-tune cell-type-specific gene expression programs.

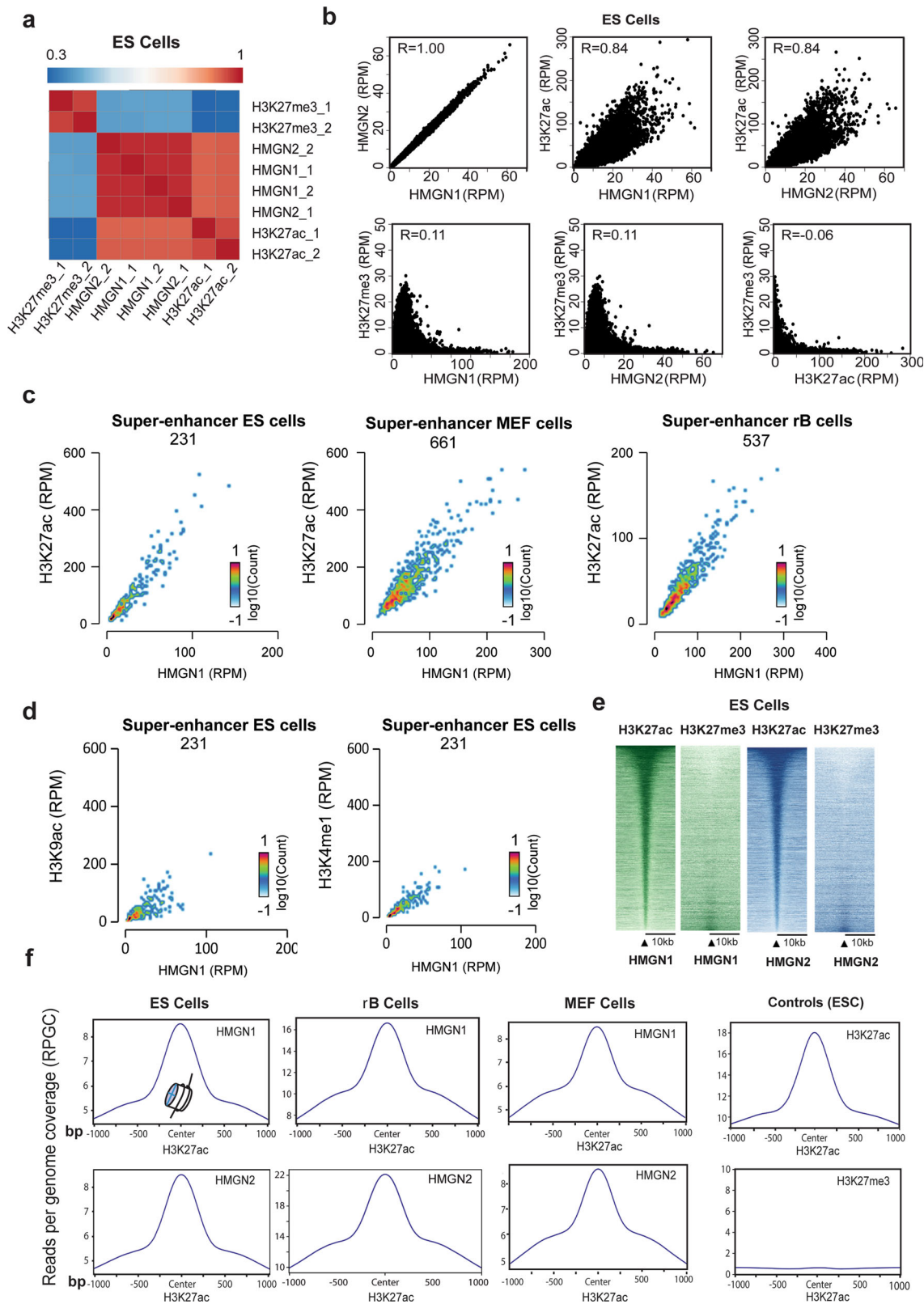
Results

HMGN1 and HMGN2 localize to H3K27ac nucleosomes.

Analyses of the genome-wide distribution of HMGN1 and HMGN2 in mouse embryonic stem cells (ESCs) (Fig. 1a, b), resting B cells (rBs) (Supplementary Fig. 1a) and embryonic fibroblasts (MEFs) (Supplementary Fig. 1b) reveal that both HMGN1 and HMGN2 variants colocalize with H3K27ac, a histone modification that marks enhancers and promoters^{7,28,29}, but not with H3K27me3, an epigenetic mark of silent chromatin. Furthermore, the intensity of both the HMGN1 and HMGN2 signal correlates positively with the intensity of the H3K27ac signal but shows no correlation with the H3K27me3 signal levels (Fig. 1b, Supplementary Fig. 1a, b). The H3K27ac signal is known to be enriched in the cell-type-specific super-enhancer regions⁷. We find that in the super-enhancer regions of ESCs, MEFs, and rB cells, the H3K27ac reads colocalizing with either HMGN1 or HMGN2 are markedly higher than the reads obtained with other epigenetic marks of active chromatin such as H3K9ac, H3K4me1, H3K64ac, or H3K122ac (Fig. 1c, d, Supplementary Fig. 1c, d, e, f). Although some of the differences in the signal intensities of the various histone modifications may be due to the quality of the antibodies used, the marked positive correlation between the signal intensity of H3K27ac and HMGNs prompted us to focus on the H3K27ac modification and investigate the factors that determine the preferential targeting of HMGN to the H3K27ac nucleosomes.

Throughout the genomes of the three cell-types, the H3K27ac signal, but not the H3K27me3 signal centers at the chromatin loci containing either HMGN1 or HMGN2 (Fig. 1e, Supplementary Fig. 1g). Significantly, high-resolution analyses of the colocalization plots reveal that the location of both HMGN1 and HMGN2 centers narrowly on the location of the H3K27ac signal, with the same precision as the H3K27ac signal itself (Fig. 1f). The overlap between the H3K27ac and the HMGN chromatin occupancy signals suggest that HMGNs localize to the H3K27ac nucleosomes (MNH3K27ac).

In sum, the ChIP seq data indicate that both HMGN1 and HMGN2 localize to nucleosomes containing the H3K27ac modification, a major epigenetic mark of cell-type-specific chromatin regulatory sites^{7,28}. These findings may provide insights into the molecular mechanism whereby HMGNs affect cell-type-specific gene expression.



Determinants that target HMGNs to H3K27ac nucleosomes.

To verify that HMGNs bind preferentially to MNH3K27ac and explore the determinants that targets these proteins to the acetylated nucleosomes, we prepared chromatin from purified MEF nuclei, stripped the chromatin binding proteins by salt extraction, digested the salt-extracted chromatin with micrococcal nuclease, fractionated the digest on a sucrose gradient, and

isolated chromatin fragments containing either only mono-nucleosomes (MN) or oligonucleosomes (ON); a mixture of tri-penta nucleosomes (Fig. 2a). To the ON fraction we added purified HMGN1 or HMGN2 at a ratio of one molecule of HMGN per 25 MNs and immunopurified the ON fraction containing bound HMGN, with antibody to HMGNs. We purified the H3 histone fraction from the input and from the immunopurified

Fig. 1 HMGN1 and HMGN2 localize to H3K27ac nucleosomes. **a** Correlation heat map showing preferential localization of both HMGN1 and HMGN2 to chromatin regions containing H3K27ac but not to regions containing H3K27me3. Bin size:5000 bp. **b** Scatter plot showing direct correlation between HMGN1 and HMGN2 occupancy levels and H3K27ac, but not H3K27me3 in ESCs. **c** Scatter plot showing direct correlation between occupancy levels of HMGN1 and H3K27ac at the super enhancers of ESCs, MEFs, and resting B cells. The number of super enhancers in each cell type is indicated below the cell name. **d** Scatter plot showing the correlation between occupancy levels of HMGN1 and either H3K9ac or H3K4me1at ESCs super enhancers. **e** Heat map showing localization of the H3K27ac and H3K27me3 signal at chromatin sites containing either HMGN1 or HMGN2. HMGN1 or HMGN2 chromatin sites were aligned at the center and the distribution of H3K27ac reads around the center were mapped. Regions are sorted in descending order based on the mean value per region. **f** High-resolution profile plots showing genome-wide co-localization of HMGNs and H3K27ac in several cell types. ES embryonic stem cells, MEF mouse embryonic fibroblasts, rB resting B cells. The source of the ChIP data for each epigenetic modification is mentioned in the Materials and Methods, in the section titled “Chromatin immunoprecipitation, Illumina library constructions and sequencing”.

ON by HPLC and performed dot-blot western analysis with antibodies to either H3K27ac, H3, H3K27me3, or H3K9ac. These dot-blot analyses reveal that the levels of H3K27ac in ON fraction that bound either HMGN1 or HMGN2 are over threefold higher than in the input ON fractions (Supplementary Fig. 2a), an indication that *in vitro*, HMGNs preferentially interact with chromatin fragments enriched in nucleosomes containing acetylated H3K27, but not with fragments enriched in either H3K27me3 or H3K9ac. Furthermore, Western analyses show that the levels of H3K27ac signal in MN-ON that bound HMGN was 2.7-fold higher than in input MNs (Fig. 2b), further indication that HMGNs bind preferentially to MNs containing H3K27ac.

HMGN proteins bind specifically to MNs and form complexes (HMGN:MN) which in a mobility shift assay, migrate slower than MN particles^{30,31} (Fig. 2c). We reasoned that the MN in complexes formed at low HMGN to MN ratio (HA, in Fig. 2c) have a higher affinity for HMGN than MNs that did not form complexes even at relatively high HMGN to MN ratios (LA in Fig. 2c). In our experiments, at an HMGN1:MN ratio of 0.2, the HA fraction contained only 9% of the total MNs, while at HMGN:MN ratio of 2.0, the LA fraction contained only 12% of the total MN (Supplementary Fig. 2b). Dot-blot westerns with H3 purified from HA and LA MNs, indicate that the levels of H3K27ac in the HA complexes are approximately 1.5 times higher than in LA MNs (Supplementary Fig. 2c, d), indicating that HMGN proteins preferentially bind to MNH3K27ac. Similar experiments using antibodies specific to either H3K27me3 or to H3K9ac did not show enrichment of modified MNs in the HA fraction (Supplementary Fig. 2c, d). In addition, Western analysis of the purified HA and LA mononucleosomes, using antibodies to either H3K27ac or H3 show that the levels of H3K27ac in the HA fraction were 2.0-fold higher than in the LA fraction (Fig. 2d, Supplementary Fig. 3), further evidence that HMGN1 preferentially bind to MNH3K27ac.

Taken together, the results show that HMGNs preferentially bind to nucleosomes and chromatin fragments containing H3K27ac but show no significant preference for nucleosomes containing other marks of active chromatin such H3K9ac and H3K4me1, or for nucleosomes containing H3K27me3, a mark of transcriptionally silent, compact chromatin.

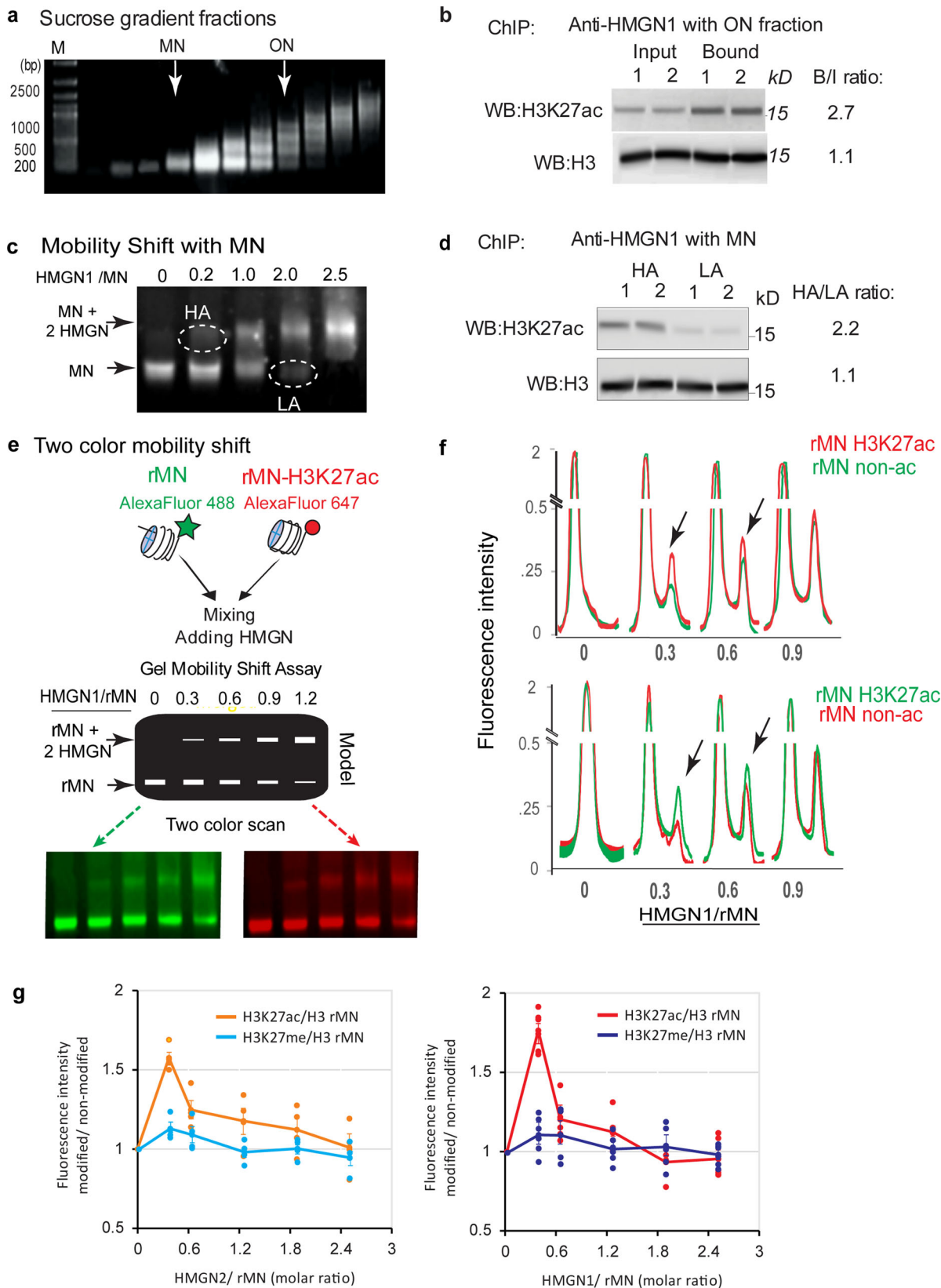
Native MNs may contain more than one histone modification³². To test whether the H3K27ac modification, by itself, is enough to preferentially target HMGN to acetylated MNs, we used commercial recombinant MN (rMN) that were devoid of any modification or contained only the H3K27ac modification (rMNH3K27ac). We end labeled the DNA in rMN with AlexaFluor 488 (green) and the DNA in rMNH3K27ac with AlexaFluor 647 (red) respectively, mixed equal amounts of the red and green fluorescent-labelled particles, performed mobility shift assays at increasing molar ratio of HMGN to the fluorescently labeled rMNs, and scanned the gels to visualize either the green or red signal (Fig. 2e). To exclude possible effects of the fluorescent label, we reversed the label and repeated the mobility shift assays with red-labeled rMN and green-labeled rMNH3K27ac.

Quantification of fluorescence scans of the gels (Fig. 2f) indicate that at low HMGN to nucleosome ratio, the slower moving fraction, which contains HMGN bound nucleosomes (rMN + 2HMGN in Fig. 2e), was enriched in rMNH3K27ac (Fig. 2f). As an additional control, we performed the same type of experiments with rMN and rMNH3K27me3 that were labeled with the same fluorochromes. Scans of these gels indicate that HMGNs do not show preferential binding to the H3K27me3 recombinant nucleosomes (Supplementary Fig. 4a, b, c, d, Supplementary Data 1). Quantitative analyses of the scans of the two-color mobility shift assays show that the ratio of the fluorescence intensity of rMNH3K27ac to rMN was as high as 1.7 while that of H3K27me3/rMN was close to 1.0 (Fig. 2g).

To further verify that *in vivo* the binding of HMGN correlates with H3K27ac level, we cultured MEF cells with the p300/CBP HAT inhibitor A-485 for 3 hours resulting in a 60% global reduction in the H3K27ac (Supplementary Fig. 5a). Subsequently, we performed HMGN1 ChIP with both A-485 treated cells and with control, untreated cells and used qPCR to determine the HMGN1 occupancy at 5 selected regulatory sites with high H3K27 modification levels and 4 sites located in gene desert regions and showing low H3K27ac levels (location of sites listed Supplementary Methods in Supplementary files). We find that A-485 treatment led to a reduction in HMGN chromatin binding at all sites (Supplementary Fig. 5b, Supplementary Data 1); at regulatory sites with high H3K27ac levels the HMGN1 occupancy were decreased by over 50% while at non-regulatory the decrease in HMGN1 was by 30% (Supplementary Fig. 5c). Thus, in living cells, downregulation of H3K27ac levels decreases the chromatin binding of HMGNs, especially at sites that show high H3K27ac reads.

In sum, ChIP-Western analysis with purified ONs, or with HMGN-MN complexes purified by gel mobility, two-color mobility shift analysis with rMN and rMNH3K27ac (Supplementary Table 1), and ChIP qPCR analysis of cells treated with an acetylase inhibitor indicate that HMGNs preferentially bind to MNs containing H3K27ac.

Next, we tested whether the preferential binding of HMGN to MNH3K27ac depends on specific regions present in HMGN proteins or on specific properties of the acetylated MNs. Two-color mobility shift assays indicate that the preferential binding of HMGN to MNH3K27ac is maintained even in the presence of 1000 fold molar excess of a competing acetylated peptide (KAARK(27ac)SAPATGG) spanning the H3K27ac residue (Supplementary Figs. 6a, 7a), an indication that the acetyl moiety, by itself, is not a major determinant of HMGN binding. HMGN proteins contain 3 functional domains: a bipartite nuclear localization signal, a highly conserved nucleosome binding domain (NBD) and a regulatory domain located in the C-terminal of the protein (Supplementary Figs. 6b, 7b)^{13,33}. It is known that HMGN deletion mutants lacking the C-terminal regulatory domain bind to nucleosomes, but even a single mutation in the NBD abolishes the binding of HMGN to MN regardless of acetylation status³⁴. We find that HMGN1 lacking the C-terminal regulatory domain still preferentially binds to rMNH3K27ac



(Supplementary Fig. 6b, right) suggesting that this domain does not determine the preference for the modified nucleosome. Thus, HMGNs do not contain distinct domains that specifically recognize the rMNH3K27ac.

The unstructured N-terminal of histone H3 interacts with nucleosome linker DNA, and HMGNs can modify the interaction

of the H3 tail with the linker DNA¹⁸. We digested the 165 bp rMN particle with micrococcal nuclease to generate the linkerless 147 bp core particles (rCP) (Supplementary Figs. 6c, 7c). Two-color mobility shift assays reveal that removal of the DNA linker region abolished the preferential binding of HMGN to rMNH3K27ac particles (Supplementary Fig. 6c, right panel).

Fig. 2 Preferential binding of HMGN to chromatin particles containing H3K27ac. **a** Agarose gel showing sucrose gradient fractionated salt stripped MEF chromatin particles. MN: mononucleosomes, ON: Oligonucleosome (mostly tri-penta nucleosomes). **b** Western analysis of total ON (Input) and HMGN1 immunoprecipitated ON (bound) **(c)** Gel mobility shift -assay. Purified HMGN1 was added to salt stripped chicken erythrocyte MNs at the ratio indicated on top of each column. The MNs shifted at low HMGN1:MN were designated as high affinity (HA) while the MNs not shifted at high HMGN1:MN were designated as low affinity (LA). **d** Western analysis of HA and LA mononucleosomes (MN). **e** Two color gel mobility shift assays of recombinant mononucleosomes (rMN). A mix of equal amounts of fluorescently Alexa 488 labeled rMN (green) and Alexa 647 labeled rMNH3K27ac (red) were incubated with various amounts of HMGN1, the mixture fractionated on native polyacrylamide gels, and the gels scanned to visualize and quantify either the red or green fluorescence. Shown is the experimental design and gel images visualized with red or green channels **(f)** Scan of the gels shown in **e** and of a similar gel in which the fluorescent labels are reversed. Top: Alexa 647 labeled rMNH3K27ac (red) and Alexa 488 labeled rMN (green). Bottom: Alexa 488 labeled rMNH3K27ac and Alexa 647 labeled rMN. Arrows point to preferential binding of HMGN to rMNH3K27ac in the shifted nucleosome. **(g)** Quantification of the scans shown in panel F and in Supplementary Fig. 4 for HMGN1 and of similar experiments done with HMGN2. Note that at low ratio of HMGN to nucleosomes, both HMGN1 and HMGN2 preferentially bind to rMNH3K27ac but not to rMNH3K27me3. Error bars represent corrected standard deviation, $n \geq 3$.

Thus, linker DNA is required for the preferential binding of HMGN to the rMNH3K27ac.

In sum, the acetylated H3K27 residue is not the major HMGN1 binding site, and HMGN proteins do not contain specific domains that lead to preferential binding to acetylated MNs. Thus, the unique properties of the MNH3K27ac, as compared to non-acetylated MNs, are the major facilitators of the preferential binding of HMGN to the modified MNs.

HMGNs modulate the levels of H3K27 modifications and histone H1 binding at chromatin regulatory sites. Next, we tested whether HMGNs affect the levels of epigenetic modifications at H3K27 residues in living cells and performed ChIP analyses with MEFs derived either from WT or from double knock-out mice lacking both HMGN1 and HMGN2 (DKO mice)²¹. Box plot analysis of these ChIP show decreased levels of H3K27ac, but increased levels of H3K27me3 at both enhancers and promoters of DKO cells (Fig. 3a). In agreement, the MA plot show numerous sites, at both enhancers and promoters, where the differences between WT and DKO cells in H3K27ac levels are statistically significant (Fig. 3b, top). Aggregate plots show changes in the global distribution of the H3K27ac only at enhancer regions, where loss of HMGNs leads to a detectable decrease in the H3K27ac occupancy levels (Fig. 3c, top lane). For H3K27me3, the MA plots show fewer statistically significant differences between WT and DKO cells in the modification level at a particular site (Fig. 3b, bottom) while aggregate plots indicate that loss of HMGNs leads to a detectable increase in the overall H3K27me3 signal at both enhancers and promoters (Fig. 3c, middle lane). Thus, loss of HMGNs decreased the level of H3K27ac but increased the levels of H3K27me3, an indication that HMGNs affect the levels of epigenetic marks at chromatin regulatory sites.

H3K27me3 levels correlate directly with the chromatin occupancy of histone H1^{8,35}, an abundant protein family that binds dynamically to nucleosomes³⁶ and stabilizes chromatin compaction⁸. Previous ChIP analyses indicated that H1 is evenly distributed throughout the nucleus but shows low chromatin occupancy at TSS^{37,38}. Our ChIP analyses do not show statistically significant differences between WT and DKO in H1 occupancy at a particular site; however, the aggregate plots indicate a marginally higher occupancy of H1 at enhancers and promoters in DKO, as compared to WT (Fig. 3c, bottom lane). Thus, the presence of HMGN lowers the H1 chromatin occupancy at regulatory sites, a finding that agrees with several previous analyses showing that HMGNs destabilize the binding of H1 to nucleosomes^{17,18} and reduce its chromatin residence time¹⁶.

At transcription start sites, the occupancy levels of both H3K27ac and HMGN show a direct correlation with gene

expression levels^{22,39}. We ranked the genes into five categories according to their expression levels and noted that indeed, the most highly expressed genes show the highest level of H3K27ac modification and the highest level of HMGN occupancy (Supplementary Fig. 8a). Loss of HMGNs increases the H3K27me3 levels and H1 occupancy at the TSS of highly expressed genes to a markedly larger degree than at low expressing genes (Supplementary Fig. 8b, c), a finding that agrees with the preferential occupancy of HMGNs on MNH3K27ac. In sum, loss of HMGN leads to increased H3K27me3 levels and elevated H1 occupancy at enhancers and promoters.

HMGNs affect the binding of regulatory factors to chromatin.

Enhancers and promoters serve as major binding sites for transcription factors (TFs). The preferential location of HMGNs at these sites and the epigenetic changes observed in DKO MEFs, together with the known effect of HMGNs on cellular transcription profiles⁴⁰ and on the stability of cell identity²², raises the possibility that HMGNs affect the interaction of TFs with chromatin regulatory sites. Therefore, we examined the effect of HMGNs on the chromatin occupancy of the acetyltransferase p300 and of the bromodomain-containing protein Brd3, a “writer” and “reader” of H3K27ac. RNA-seq and western blot analyses of extracts prepared from WT and DKO MEFs, indicate that loss of HMGN did not affect the transcript and protein levels of either p300 (Fig. 4a, Supplementary Fig. 9a) or Brd3 (Fig. 4e, Supplementary Fig. 9b). Yet, ChIP analyses reveal that loss of HMGNs leads to a marked decrease in the chromatin occupancy of both p300 (Fig. 4b, c, d, i) and of Brd3 (Fig. 4f, g, h, i) throughout the MEF genome and at both enhancers and promoters. The number of significantly decreased p300 binding sites at enhancers and promoters, (869 and 427, respectively) was 20-fold higher than the sites that show increased occupancy (Fig. 4b, c). The chromatin binding of Brd3 was affected to a larger degree; at enhancers loss of HMGNs decreased significantly Brd3 binding at 5862 sites but increased the Brd3 binding at only 64 sites (Fig. 4g). Similar effects are seen at MEF promoters where the loss of HMGN decreased Brd3 binding at 5090 sites (Fig. 4g). In agreement, aggregate plots show decrease chromatin occupancy of p300 and Brd3 at both enhancers and promoters (Fig. 4d, h).

Likewise, the expression levels of CEBPB, a TF that binds to chromatin with DNA sequence specificity⁴¹, were not affected by the loss of HMGN (Fig. 5a, Supplementary Fig. 10), but the CEBPB chromatin occupancy in DKO MEFs is noticeably diminished throughout the genome (Fig. 5b) and at both promoters and enhancers (Fig. 5c). Of the 19,392 CEBPB sites detected in WT cells, 10,865 and 8527 sites were lost and retained, respectively in DKO MEFs (Fig. 5d). The top CEBPB binding sequence motif in both WT and DKO cells corresponds to the canonical CEBPB binding motif (Fig. 5e), an indication that

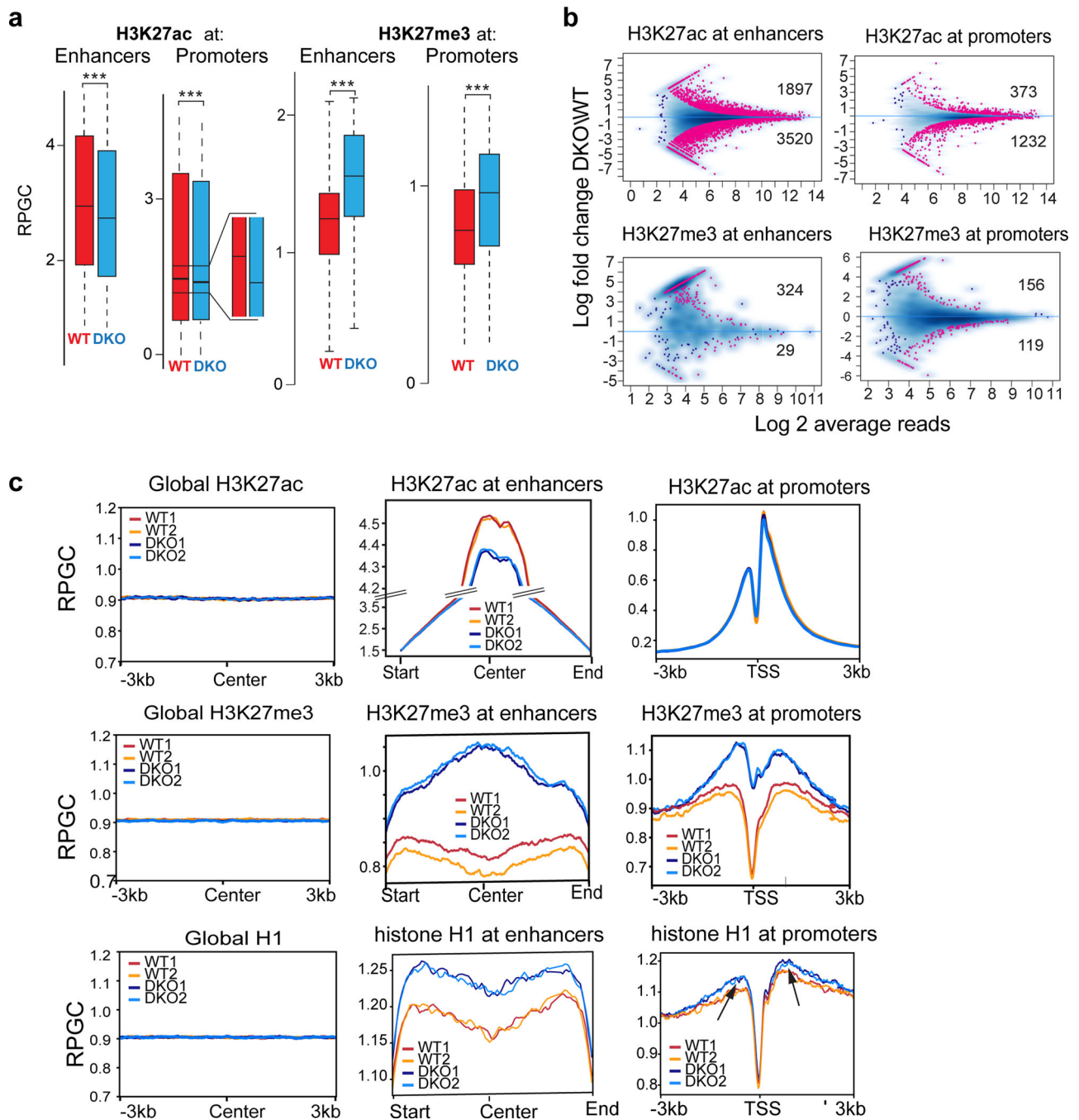


Fig. 3 Loss of HMGN alters H3K27 modifications and H1 occupancy at chromatin regulatory sites of MEFs. **a** Box plot showing decreased H3K27ac but increased H3K27me3 levels at enhancers and promoters of DKO mice. **b** MA plots showing differences between WT and DKO cells in H3K27ac or H3K27me3 levels at enhancers and promoters. Statistically significant differences (FDR < 0.05) are shown in red. Blue dots and blue density cloud represents all points corresponding to the nonchanging regions. **c** Aggregate plots showing the distribution of the average H3K27ac, H3K27me3 and histone H1 levels in WT and DKO MEFs. Left panels: throughout the genome (regulatory sites subtracted). “Center” indicates a location of the middle point of each 6kbp bin. Center panels: at enhancers. In these panels all cellular enhancers were aligned at their center. RPGC: reads per genomic coverage. Right panels: at promoters, Arrows point to promoter regions where H1 occupancy differs between WT and DKO cells. All ChIP analyses from two biological replicates. Regulatory sites identification is based on UCSC genome annotation (NCBI37/mm9).

HMGN does not affect the DNA binding sequence specificity of CEBPB. For both lost and retained sites, the CEBPB ChIP seq peaks align narrowly on the center of the CEBPB binding motif, (Fig. 5f, top two lines); however, a search for TF DNA-binding sequence motifs uniquely present only in retained, or only in lost CEBPB sites show differences between these sites. The retained CEBPB sites are flanked by DNA binding motifs for additional transcription factors such as FOSB, JUN, and ATF3 (Fig. 5f, third

line), while the lost CEBPB sites are not surrounded by known transcription factor binding motifs (Fig. 5f bottom line). In addition, the H3K27ac levels at retained sites are higher than at the lost sites (Fig. 5g). In WT MEFs, 64% of CEBPB binding sites localized to chromatin regions showing H3K27ac reads; 86% of these also showed HMGN1 and HMGN2 occupancy. At nonacetylated sites, the co-occupancy of CEBPB with HMGN was only 39% (Fig. 5h). Representative IGV screenshots

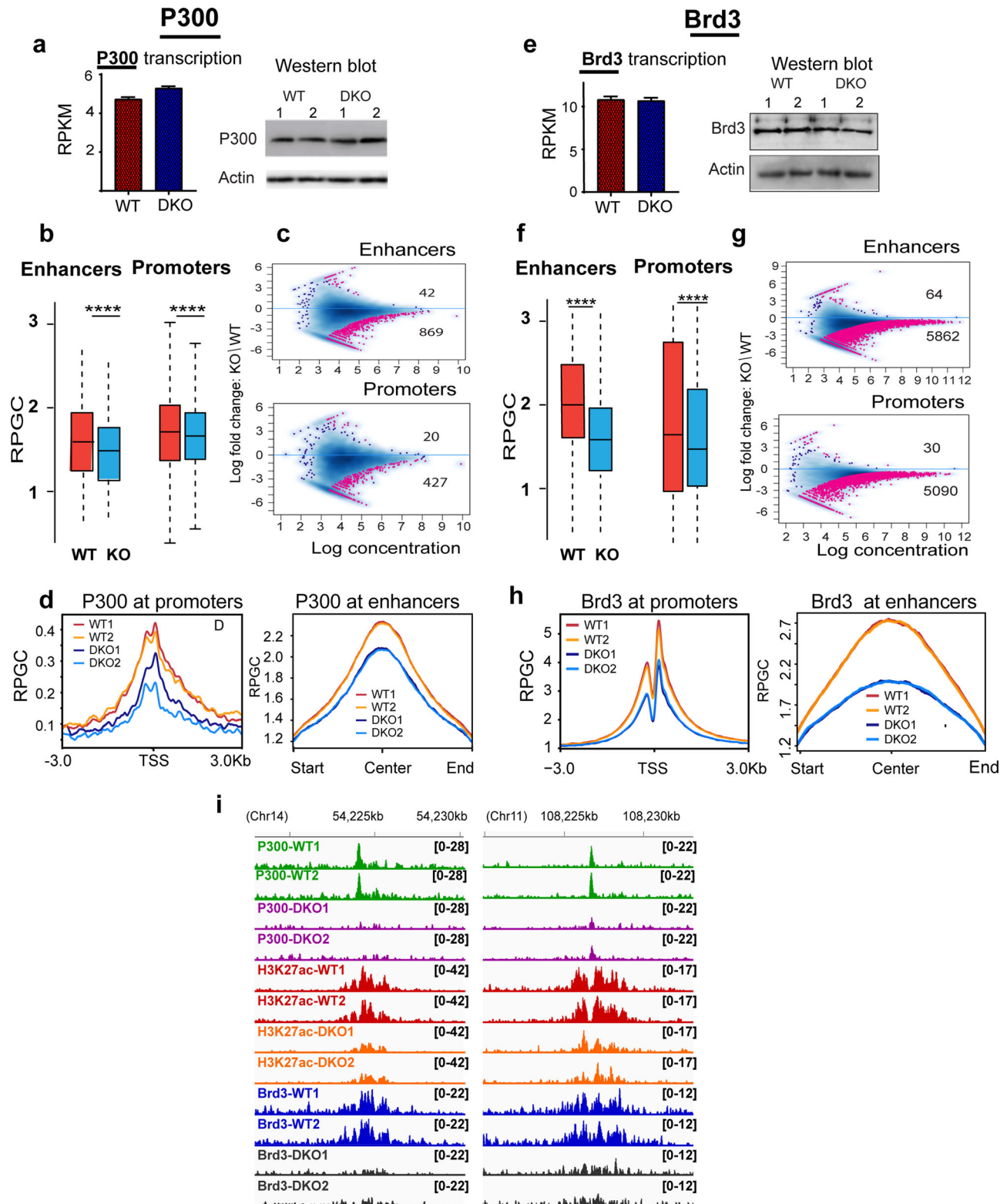


Fig. 4 Decreased p300 and Brd3 chromatin binding in DKO MEFs. **a** Equal P300 expression in WT and DKO MEFs. **b** Box plots showing decrease P300 chromatin occupancy at enhancers and promoters of DKO cells. **c** MA plots showing differences in P300 chromatin binding between DKO and WT cells. Sites showing statistically significant differences (FDR < 0.05) are in red. Blue dots and blue density cloud represents all points corresponding to the non-changing regions. Note that most altered sites show decrease binding in DKO cells. **d** Profile plots showing decreased P300 occupancy at promoters and enhancers of DKO cells. **e** Equal Brd3 expression in WT and DKO MEFs. **f** Box plots showing decrease Brd3 chromatin occupancy at enhancer and promoters of DKO MEFs. **g** MA plot showing differences in Brd3 chromatin binding between DKO and WT cells. Sites showing significant differences (FDR < 0.05) are in red. Blue dots and blue density cloud represents all points corresponding to the non-changing regions. **h** Profile plots showing decreased Brd3 occupancy at promoters and enhancers of DKO cells. **i** IGV tracks showing reduced H3K27ac levels, and reduced P300 or Brd3 chromatin occupancy in DKO cells. All CHIP analyses from two biological replicates.

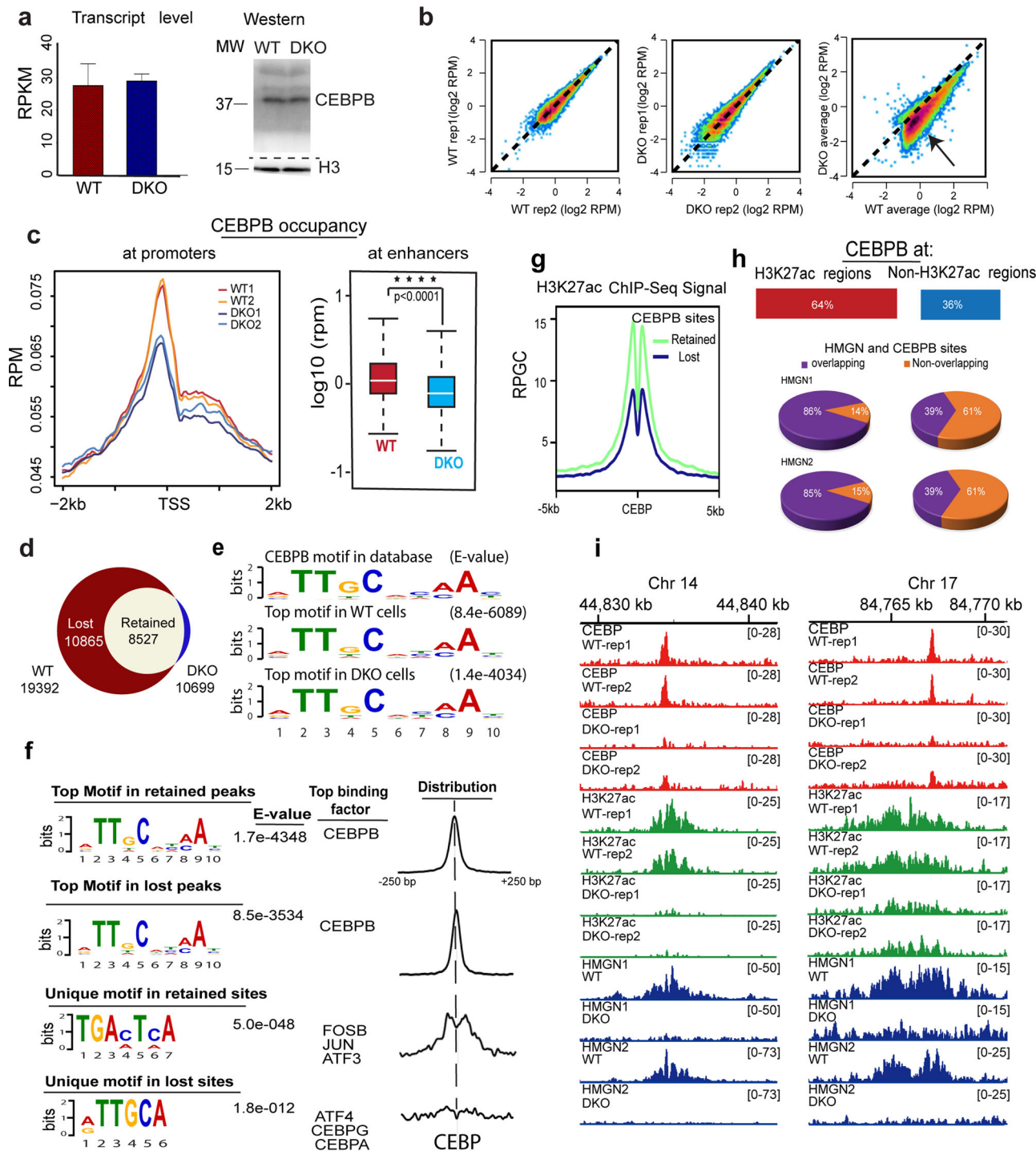


Fig. 5 Decreased CEBPB chromatin binding in DKO MEFs. **a** Equal levels of CEBPB transcript and protein in WT and DKO MEFs. **b** Scatter plot comparing intensities of CEBPB peaks between biological replicates of WT (left), and of DKO cells (center). Right scatter plot shows reduced CEBPB chromatin binding in DKO cells. **c** Decreased CEBPB binding at TSS and enhancers of DKO cells. **d** Venn diagram showing CEBPB chromatin binding sites in WT and DKO MEFs. **e** Top DNA sequence motif underlying the CEBPB binding sites in WT and DKO cells, compared with the CEBPB motif in database. **f** Top and unique motifs in retained and lost CEBPB binding sites. Lost CEBP sites are defined as present in WT but not in DKO cells. The diagrams to the right show the location of the DNA binding motifs relative to the center of the CEBPB binding motif. **g** H3K27ac levels at lost or retained CEBPB sites in WT cells. **h** Overlap between CEBPB, H3K27ac and HMGN occupancy. **i** IGV snapshots showing loss of CEBPB binding in DKO cells at regions overlapping with H3K27ac. All ChIP analyses from two biological replicates.

visualizing epigenetic changes at CEBPB binding sites are shown in Fig. 5i.

Thus, the binding of CEBPB to chromatin is modulated by, but not exclusively dependent on, the presence of HMGN protein. CEBPB binding sites that contain motifs for additional

transcription factors or show high H3K27ac reads are less affected by the loss of HMGNs than sites that show low H3K27ac levels and are not in proximity to additional TFs. Similar studies show that HMGNs affect the chromatin interactions of TFs known to play a role in the development and function of mouse B

cells. ChIP analyses of IKAROS, ETS1, IRF8 and PAX5 in resting B cells isolated from WT and DKO mice show that invariably, loss of HMGN reduced the binding of the TFs to chromatin throughout the genome and at both enhancers and promoters (Supplementary Fig. 11).

In ESCs, loss of HMGN leads to a marked reduction in the chromatin binding of the H3K27ac reader Brd4 (Fig. 6, top lane), and to a more moderate loss of chromatin binding of Klf4 (Fig. 6, 2nd lane) and CTCF (Fig. 6, 3rd lane), two TFs known to affect global chromatin topology⁴². Interestingly, the pluripotency factors NANOG, SOX2 and OCT4, whose nucleosome binding motifs are located in proximity to HMGN binding sites, show higher chromatin occupancy in DKO ESCs than in WT cells (Fig. 6, bottom 3 lanes).

Thus, ChIP analyses in MEFs, rBs and ESCs consistently show that loss of HMGN alters the interaction of TFs with chromatin; however, the magnitude and type of effect is context-dependent on the exact mode of interaction of a TF with its cognate binding site in chromatin.

Discussion

H3K27ac is a major epigenetic mark of active chromatin while HMGNs are ubiquitous nuclear proteins that bind dynamically to chromatin without specificity for DNA sequence. Previous studies demonstrated that genome wide, HMGNs colocalize with H3K27ac^{21,22}. Super-enhancers, regions known to have a high content of H3K27ac nucleosomes, show a strong correlation between HMGN occupancy and cell-type-specific super-enhancer activity; HMGNs localize to a specific genomic region only when it acts as a super-enhancer, i.e., has a high content of H3K27ac nucleosomes. The same sequence (in another cell type, or at different differentiation stage) that does not serve as a cell-type-specific super enhancer, and does not have a high content of H3K27ac, does not show high HMGN occupancy²². Furthermore, during reprogramming of MEFs to induced pluripotent stem cells (iPSC), HMGNs relocate from the acetylated MEFs enhancers regions to the acetylated iPSCs regulatory regions, supporting the notion that HMGN bind preferentially to the regions containing H3K27ac regardless of the underlying DNA sequence²². Now we show a direct correlation between the signal intensity of H3K27ac and HMGN occupancy (Fig. 1b) and using an acetylase inhibitor we show that in living cells, a decrease of H3K27ac reduces HMGN1 occupancy. Thus, multiple types of experiments show that HMGNs bind preferentially to chromatin regions containing high levels of H3K27ac nucleosomes.

Here we analyze the determinants that facilitate the binding of HMGN to regulatory regions enriched in H3K27ac nucleosomes and identify H3K27ac as an epigenetic signal that facilitates the localization of HMGNs to specific chromatin sites, thereby leading to epigenetic changes that affect the binding of TFs to chromatin. Likely, H3K27ac is not the only epigenetic mark that affects the binding of HMGN to chromatin; additional epigenetic marks may also contribute to the preferential localization of HMGN to the regulatory region. Indeed, previous analyses^{21,22} revealed that genome wide, within a span of 5000 bp or more, HMGNs colocalize with several epigenetic marks of active chromatin. Now we show colocalization of HMGN with H3K27ac nucleosome at a higher resolution, within a span of 160 bp, i.e. the span of a single nucleosomes. Our findings that H3K27ac-modified nucleosomes help recruit HMGN proteins to chromatin regulatory sites provide novel insights into epigenetic factors that fine-tune cell-type-specific transcription and stabilize cell identity.

HMGN proteins bind to nucleosomes regardless of whether they are or are not acetylated at H3K27. However, in the genome, HMGNs colocalize with H3K27ac because they bind

preferentially to H3K27 acetylated MNs, as compared to non-acetylated MNs. This preference is seen not only in the genome of cultured cells but also in experiments containing only HMGN and recombinant MNH3K27ac, an indication that the increased binding of HMGN to chromatin regions containing H3K27ac is not dependent on special cofactors, on unique nucleosome spacing, or on the presence of additional modifications on the targeted nucleosomes. Thus, the presence H3K27ac, is sufficient to preferentially target HMGNs to MNH3K27ac. The preference is not due to the acetyl moiety itself since a peptide containing the H3K27ac residue does not inhibit the binding of HMGN to MNH3K27ac, and HMGNs do not show preference for H3K9ac nucleosomes. Considering the mechanisms driving the preference for MNH3K27ac, we note that previous studies show that HMGNs bind to MNs through a conserved nucleosome binding domain that contacts the nucleosome near the nucleosome dyad⁴³ and at the nucleosome acidic patch in the H2A.H2B dimer¹⁹. Point mutations in this domain abolish the binding of HMGN to nucleosomes suggesting that this domain is not involved in the preferential binding of HMGN to acetylated nucleosomes. The HMGN C-terminal contacts the N-terminal of H3³³ and disrupts its interaction with the linker DNA¹⁸, yet we find that deletion of the HMGN C-terminal domain does not abolish the preferential binding of HMGN to MNH3K27ac. These analyses suggest that an HMGN protein does not contain specific regions that can distinguish between acetylated and non-acetylated MNs.

We identify two major factors that determine the preferential binding of HMGN to MNH3K27ac: the presence of the acetylated H3K27 residue and the presence of linker DNA, yet the HMGNs interaction with the acetylated H3K27 residue is not the major factor determining the preferential binding to acetylated MNs and HMGNs bind well to the linker-less 146 bp core particle. In considering how acetylation of H3K27 affects HMGN binding, we note that the unstructured H3 tail interacts with both the linker DNA and with the DNA surrounding the histone octamer. The interaction of the H3 tail with nucleosomal DNA can affect nucleosome dynamics^{44,45} and alter the DNA conformations, especially in regions close to the nucleosome dyad axis, a region where HMGNs bind to the nucleosome^{18,43}. Modifications such as lysine acetylation can alter the local conformation of the H3 histone tail and its interaction with the DNA⁴⁶. Thus, together with previous information, our results suggest that acetylation of H3K27 leads to conformational changes in the nucleosome that facilitate binding of HMGN, thereby increasing the time that an HMGN molecule resides at a specific chromatin regulatory site.

An important consequence of increased HMGN chromatin residence time is a decrease in the chromatin residence time of histone H1¹⁶, a protein known to promote chromatin compaction⁸. Conversely, a decrease in HMGN levels enhances the interaction of H1 with chromatin, as we previously observed in studies of specific genomic loci^{47,48} and in this study globally, using cells derived from mice lacking both HMGN1 and HMGN2 (DKO cells). These effects are most obvious at enhancers and promoters, regions which are marked by a high level of H3K27ac and high HMGN occupancy. The most highly acetylated promoters, which also show the highest HMGN occupancy in WT cells, show the highest increase in H1 occupancy in DKO cells (Supplementary Fig. 8). Histone H1 facilitates the binding of PRC2-EZH2 to chromatin and stimulate the methylation of H3K27^{8,49}. Indeed, DKO cells show increased H3K27me3 levels at enhancers and the highest increase in H3K27me3 levels at the most active promoters, i.e. promoters that in WT cells showed the highest acetylation levels and HMGN occupancy (Supplementary Fig. 8). Nevertheless, we have not excluded the possibility that HMGN also affects the binding of histone methylases to chromatin.

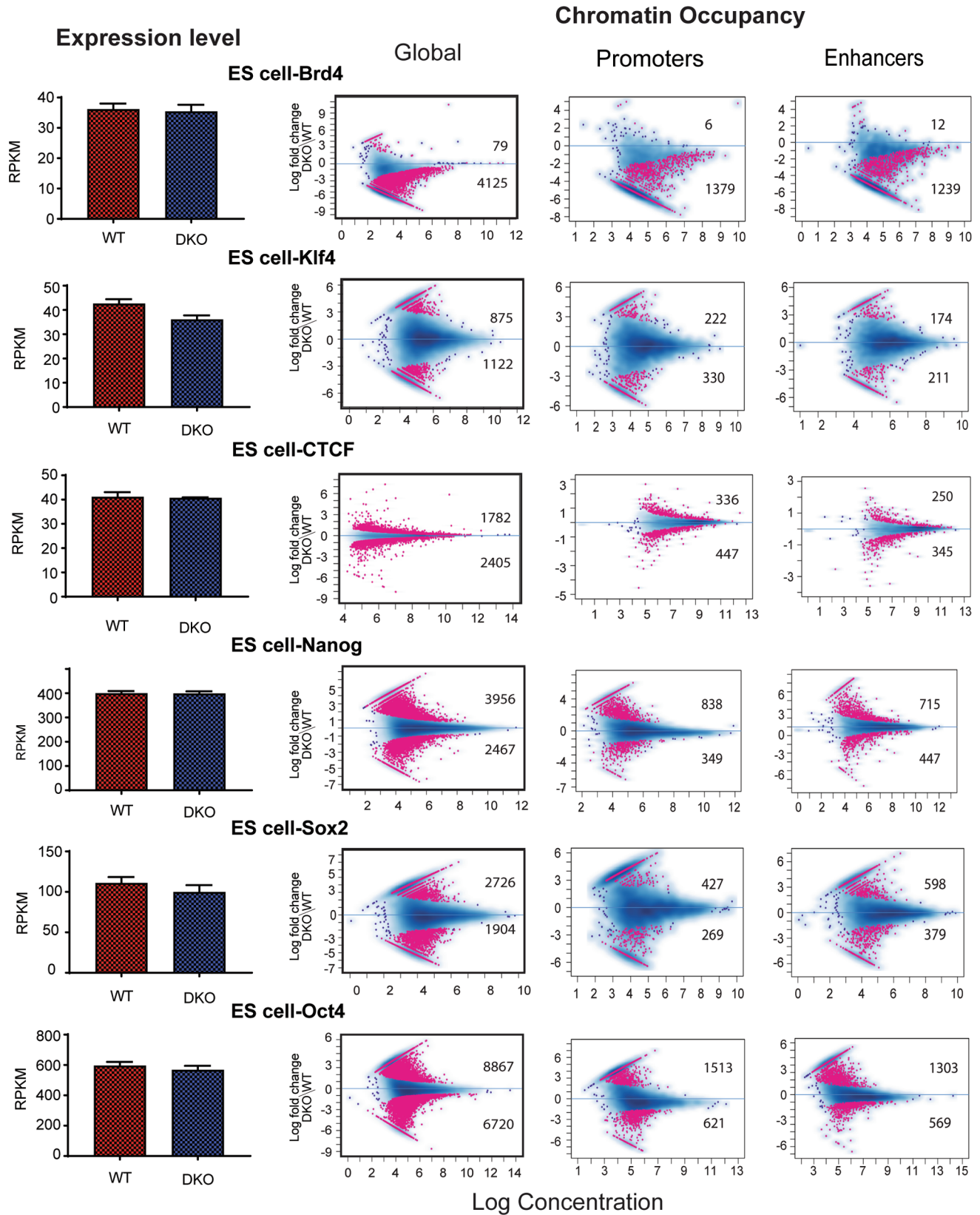


Fig. 6 Altered chromatin occupancy of transcription factors in DKO ESCs cells. Bar graphs on left show transcript levels determined by RNA seq analysis. MA plots show differences in TF chromatin binding between DKO and WT cells. Sites showing significant differences (FDR < 0.05) are in red; the number of up- and downregulated sites are indicated in each panel. Blue dots and blue density cloud represent all points corresponding to the non-changing regions. All data from 2 biological replicates.

Given the preferential binding of HMGN to MNH3K27ac in active chromatin, it could be expected that loss of HMGN would affect the binding of TFs to chromatin. Indeed, loss of HMGNs decreased the chromatin binding of most of the TFs analyzed, especially that of factors that interact with H3K27 residues such as p300, Brd3, and Brd4; their chromatin binding was markedly reduced in cell lacking HMGNs. Detailed analysis of CEBPB binding indicates that the HMGN effects depend on several additional factors, including the local levels of acetylation and the presence of cofactors that affect the binding of a specific TFs to chromatin. Interestingly, loss of HMGNs do not always reduce the binding of TFs to chromatin. We find that SOX2, OCT4 and NANOG, TFs whose binding sites^{50,51} are in proximity to the nucleosomal binding sites of HMGNs⁴³, show increased chromatin binding in DKO. The interaction of Sox2 and Oct4 with nucleosomes can lead to detachment of DNA termini from the histone octamer and to distortion in the histone-DNA contacts. The HMGN nucleosome binding sites have been mapped to the major groves flanking the nucleosome dyad axis⁴³ and thermal denaturation studies indicate that HMGNs stabilize the structure of MN by minimizing the unraveling of the DNA strands at the end of the particle^{30,52}. Thus, the increased chromatin occupancy of SOX2, NANOG, OCT4 in DKO cells agree with the known location of HMGN on the nucleosome and the effects of HMGN on nucleosome stability. Most likely, for these transcription factors, the presence of HMGN hinders access to their specific binding site in the nucleosome and may hamper their ability to unravel the structure of the nucleosome.

Taken together, the results suggest that the effects of HMGNs depend on the exact mode of TF interaction with chromatin. HMGNs did not significantly alter H3K27ac levels at promoters, perhaps because at these sites the chromatin accessibility is enhanced by multiple regulatory factors, minimizing the specific impact of HMGN. HMGNs modulate and fine-tune rather than absolutely determine their chromatin binding of TFs. Nevertheless, changes in HMGN levels do alter gene expression and destabilize cell identity, supporting the finding that HMGNs alter the binding of cell-type-specific TFs to chromatin^{22,23,53,54}.

The H3K27ac mediated recruitment of HMGN to chromatin regulatory regions provides a molecular mechanism for the experimental findings from many laboratories, which repeatedly show that changes in HMGN levels alter cell-type-specific gene expression²³. Although HMGNs bind dynamically to chromatin and constantly move through the entire nucleus, they preferentially localize to cell-type-specific super enhancers, chromatin regions that are enriched in MNH3K27ac. The relatively long residence time of HMGNs at these regulatory sites reduces the chromatin residence of H1 and facilitates TFs access to their specific sites, thereby stabilizing cell identity^{22,24,53}. Changes in HMGN levels can lead to epigenetic changes that affects the binding of TFs to their specific sites (see model on Fig. 7), resulting in cell-type-specific changes in gene expression that could affect the cellular phenotype. Indeed, mice lacking both HMGN1 and HMGN2 show multiple phenotypes²¹, reflecting the ubiquitous HMGN expression in all vertebrate cells. In humans, the increase incidents of B cells acute lymphoblastic leukemia seen in Down syndrome was directly attributed to epigenetic changes and altered transcription mediated by increased HMGN1 levels due to the extra copy of *HMGN1*, which is located on human chromosome 21^{25,54}. Significantly, in both human and mouse cells, overexpression of HMGN1 leads to upregulation of H3K27ac and downregulation of H3K27me3 levels²⁶, further evidence that altered HMGN levels can lead to epigenetic changes that affect the fidelity of cell-type-specific gene expression and impact the cellular phenotype.

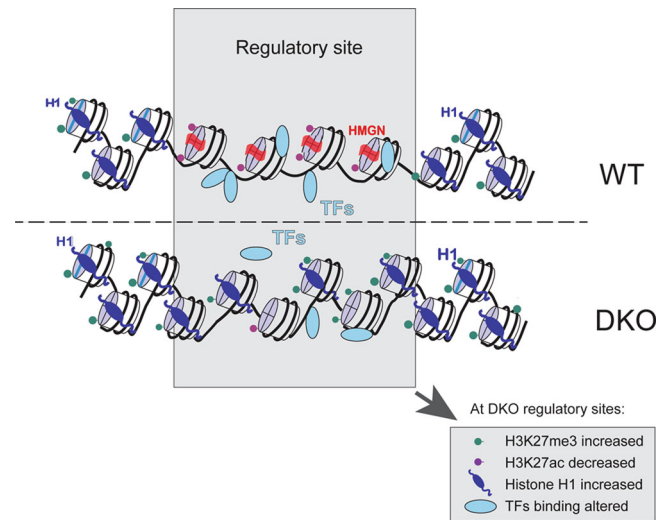


Fig. 7 Model of HMGN-mediated epigenetic changes occurring at chromatin regulatory sites. The major changes are indicated in the boxed region at the bottom of the image.

Methods

Antibodies, recombinant nucleosomes, peptides, and cell lines. Rabbit polyclonal to H1, HMGN1, HMGN2, and H3 were from our laboratory, anti H3K27ac (Abcam#ab4729), anti H3K27me3 (Abcam#ab6002), monoclonal anti H1 (Millipore-Sigma #05-457), anti CEBPB (Abcam#ab32358), Anti-Brd3 (Active Motif #61489), Anti-Brd4 (Bethyl Laboratories #A301-985A100), Anti-CEBPB (Abcam #ab32358), Anti-CTCF (EMD Millipore #07-729), Anti-Ets1 (Active Motif #39580), Anti-Ikaros (Active Motif #39355), Anti-Irf8 (Bethyl Laboratories #A304-027A), Anti-Klf4 (Abcam #106629), Anti-Nanog (Active Motif #61419), Anti-Oct4 (Abcam #ab19857), Anti-p300 (Active Motif #61401), Anti-Pax5 (Abcam #183575), Anti-Sox2 (Abcam #97959).

The following recombinant mononucleosomes were purchased from Active Motif: unmodified (#81070); H3K27me3 modified (#81834), H3K27ac modified (#81077).

Wild type and HMGN DKO mouse embryonic fibroblasts, embryonic stem cell lines⁵⁵ and resting B cells⁵⁶ were as previously described. Peptides Histone H3 (23–34) peptide, KAARKSAPATGG and Histone H3K27ac (23–34) peptide, KAAR - K(Ac)—SAPATGG were from AnaSpec, Inc.

ChIP-Western for HMGN1 and HMGN2. Mono and oligo nucleosomes devoid of protein (salt stripped chromatin particles) were prepared from mouse embryonic fibroblasts (MEFs) derived in our laboratory and from chicken erythrocytes (Rockland R401-0050) as described^{57,58}. Briefly, purified nuclei were digested by micrococcal nuclease, the chromatin prepared, non-core histone proteins removed by centrifugation in 0.5 M NaCl solutions and cation-exchange chromatography, and the salt-stripped chromatin particles loaded on 5–20% sucrose gradient. Fractions containing either mononucleosomes (MN) and oligonucleosomes (ON, containing 2–5 nucleosomes) were pooled and dialyzed against 10 mM NaCl, 10 mM Tris-Cl, pH 7.5, 1 mM EDTA. For ChIP-Western analysis the dialyzed nucleosomes were mixed with recombinant HMGN1 or HMGN2 (HMGN:histoneH3 = 50:1 molar ratio) and cross-linked with 1% formaldehyde for 10 minutes at room temperature. After quenching with 0.5 M glycine the HMGN-nucleosome complexes were immunoprecipitated using ChIP-IT Express kit (Active Motif, Cat. No. 53008) with either anti-HMGN1, anti-HMGN2 antibodies, or normal rabbit IgG as control. The immunoprecipitated nucleosomes and 2% of the input material were de-cross-linked by heating to 95° for 45 min, and histone H3 isolated by HPLC (Agilent Technologies 1200 series) using a Luna 5 μm CN 100 Å HPLC column (Phenomenex, Cat. No. 00D-4255-E0). Equal amounts of H3, as determined by OD220, from bound and input material were blotted on Immobilon PVDF membranes (Millipore, Cat. IPVH304F0), using Schleicher & Schuell Minifold Spot-Blot System and the blots subjected to western analyses.

Fluorescent labelling of mononucleosomes. For in vitro binding studies, commercial recombinant mononucleosomes (rMN), or rMN containing H3K27 modifications (Active Motif) were first end-labelled with aminoallyl dUTP using recombinant Terminal Deoxynucleotidyl Transferase (TdT), and then concentrated by spin-dialysis (Ultracel 30 K, Millipore). The dialyzed rMNs were then labeled with either Alexa Fluor 488 (Green-Fluorescent) or Alexa Fluor 647 (Red-Fluorescent) with succinimidyl ester labeling kits (ARES DNA Labeling Kit, Invitrogen, Cat A21665 and A21676). The fluorescent labelled rMN were concentrated by spin-dialysis.

Two-color gel mobility shift assays. Mobility shift assays were performed as described⁵⁹ at ionic strength conditions that lead to binding of 2 molecules of HMGN per nucleosome. Binding reactions contained 100 nM of chromatin particles (25 ng/μl) in 10 μl binding buffer (2×TBE, 0.15 mg/ml BSA and 5% Ficoll). HMGN1, HMGN2 proteins were added to chromatin particles to generate the molar ratio to nucleosomes listed in the figure legends. The mixtures were incubated at 4 °C for 10 min and loaded onto non-denaturing 5% polyacrylamide gels in 2× TBE (45 mM Tris-Borate, pH 8.3, 1 mM EDTA) and run at 4 °C. The gel was scanned with ChemiDoc MP (BioRad) using duplex fluorescence detection mode (for Alexa Fluor 488 and Alexa Fluor 647). The images were analyzed by Image Lab Touch Software and QuantityOne (BioRad). The ratios between nonshifted nucleosomes (i.e., devoid of HMGN) and shifted (HMGN-bound nucleosomes) were calculated for every titration point.

Chromatin immunoprecipitation, Illumina library construction and sequencing.

The ChIP-Seq procedure was performed as recommended by Active Motif (Carlsbad, CA) Instruction Manuals for ChIP-IT High Sensitivity, ChIP-IT Express, and Chromatin IP purification kits. Briefly, about 10⁷ cultured cells were fixed in medium with 1% formaldehyde (v/v) for 10 min at room temperature on a rocking platform, followed by quenching with 125 mM glycine. Crosslinked cells were washed and incubated in 1 ml Chromatin Prep Buffer containing 1 μl proteinase inhibitor cocktail (PIC) and 1 μl of 100 mM PMSF for 10 min on ice followed by centrifugation at 1250×g for 3 min at 4 °C. The pellets were resuspended in 250 μl ChIP Buffer with 2.5 μl PIC and 2.5 μl 100 mM PMSF and sonicated for 10 cycles with Bioruptor (30 s on/ 30 s off). Aliquots of 25 μl of sonicated chromatin were used to generate the input DNA. 5–10 μg of affinity-pure ChIP-grade antibodies were then added to the rest of the chromatin samples and incubated overnight at 4 °C with rotation. Following incubation 30 μl of protein G agarose beads or Magnetic Beads were added to each reaction and the mixtures were further incubated for 3 h or O/N at 4 °C. The beads were washed five times with Wash Buffer AM1 (Active Motif). ChIP DNA was eluted in 100 μl Elution Buffer AM4 (Active buffer). Cross-links were reversed at 65 °C overnight in the presence of 3 μl of 10% SDS and 5 μl of proteinase K (20 mg/ml). The DNA samples were eluted in 21 μl of elution buffer using MiniElute kit (Qiagen). ChIP-seq library was prepared following the manufacturer's instructions (Illumina). Briefly, immunoprecipitated and input DNA were blunt ended, ligated to adapters, amplified with PCR and size selected. The ChIP templates were sequenced at 75 bp single read length with Illumina NextSeq 500 system or 101 bp paired-end length with HiSeq 2000 by the NIH CCR sequencing facility (for details see data submission file). For the various samples the number of trimmed reads, successfully mapped to the mouse genome, ranged from 12 to 74 million per sample, with an average of 27 million reads with over 80% of trimmed, non-duplicated reads mapped to the genome. Sequence reads were aligned to the Build 37 assembly of the National Center for Biotechnology Information mouse genome data (NCBI37/mm9). Super-enhancers and enhancers were identified as described before (21). Data for the following histone marks and HMGN were downloaded from GEO archive (accession numbers are indicated in brackets) and processed in the same way as described in the methods section: ES cells: H3K9ac (GSM2417092), H3K4me1 (GSM2629668), H3K27ac (SRP154652) H3K122ac (SRR3144856), HMGNs (SRP154652), H3K27me3 (SRP068453);MEF cells: H3K9ac (GSM1979773), H3K4me1 (GSM3272827), HMGNs(SRP154652); rB cells-H3K27ac(SRP154652), HMGNs(SRP154652). H3K27me3(GSM2184272) All of the rest of the data, were generated in our laboratory.

Quantification and Statistical Analysis. Chromatin binding peaks were initially selected for analysis based on a q-value cutoff 0.01 for broad and 0.05 for narrow regions as reported by the MACS2 peak-calling algorithm (broad or narrow). Peaks were identified as significantly differentially bound using the default threshold of FDR < 0.05. Differentially expressed genes between WT and DKO cells were binned according to the average expression of 3 WT and DKO samples, with 2794 genes assigned to each group. Statistical analyses were performed within the R (ver. 3.6) computing environment and visualized with IGV. Details of statistical analyses can be found in figure legends.

Adapter trimming was performed using CutAdapt v.1.16. Sequence quality before and after trimming was checked with FastQC 0.11.5 tool. Sequences were checked for contamination with Kraken v1.1 (<https://doi.org/10.1186/gb-2014-15-3-r46>) and FastQscreen v.0.9.3 (https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen) applications. Reads were mapped to UCSC mm9 reference genome using (<https://doi.org/10.1093/bioinformatics/btp324>). Duplicate reads were marked with Picard v. 2.17.11. Reads mapped to blacklisted regions (<https://doi.org/10.1038/nature11247>) were removed from further analysis. For all samples, peak detection of enriched binding regions was performed using either MACS2 v. 2.1.1.20160309 with the default settings, or SICER v. 1.1 with the following parameters: window size 300, gap size 600, FDR < 1e-2, effective genome size 0.75. BigWig files were used for visualization. Correlation heatmaps showing sample relations, scatterplots and peak profiles were generated using DeepTools v. 3.0.1 toolset. Differential binding sites were identified using DiffBind package. To calculate Pearson or Spearman correlation coefficients between different histone mark profiles, the genome was split into bins (bin size =5000 bp) followed by

counting numbers of aligned reads within each bin for each sample and calculating a correlation between these sets.

RNA expression levels were determined by RNA-seq as described^{21,48,56}. Super Enhancer regions were downloaded from SuperDB Database as described²².

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All sequencing data generated as a part of this study are deposited to NCBI and are available with the accession number: GSE156697. Source data can be found in Supplementary Data 1. All other data are available from the authors on reasonable request.

Received: 27 May 2021; Accepted: 1 February 2022;

Published online: 23 February 2022

References

- Lambert, S. A. et al. The Human Transcription Factors. *Cell* **172**, 650–665 (2018).
- D'Alessio, A. C. et al. A Systematic Approach to Identify Candidate Transcription Factors that Control Cell Identity. *Stem Cell Rep.* **5**, 763–775 (2015).
- Voss, T. C. & Hager, G. L. Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nat. Rev. Genet.* **15**, 69–81 (2014).
- Phair, R. D. et al. Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. *Mol. Cell Biol.* **24**, 6393–6402 (2004).
- Zaret, K. S., Lerner, J. & Iwafuchi-Doi, M. Chromatin Scanning by Dynamic Binding of Pioneer Factors. *Mol. Cell* **62**, 665–667 (2016).
- Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* **20**, 207–220 (2019).
- Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
- Fyodorov, D. V., Zhou, B. R., Skultchi, A. I. & Bai, Y. Emerging roles of linker histones in regulating chromatin structure and function. *Nat. Rev. Mol. Cell Biol.* **19**, 192–206 (2018).
- Hergeth, S. P. & Schneider, R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep.* **16**, 1439–1453 (2015).
- Jordan, A. Histone H1 in gene expression and development. *Biochim Biophys. Acta* **1859**, 429–430 (2016).
- Bustin, M. High mobility group proteins. *Biochim Biophys. Acta* **1799**, 1–2 (2010).
- Reeves, R. Nuclear functions of the HMG proteins. *Biochim Biophys. Acta* **1799**, 3–14 (2010).
- Postnikov, Y. & Bustin, M. Regulation of chromatin structure and function by HMGN proteins. *Biochim Biophys. Acta* **1799**, 62–68 (2010).
- Gonzalez-Romero, R., Eirin-Lopez, J. M. & Ausio, J. Evolution of high mobility group nucleosome-binding proteins and its implications for vertebrate chromatin specialization. *Mol. Biol. Evol.* **32**, 121–131 (2015).
- Catez, F. & Hock, R. Binding and interplay of HMG proteins on chromatin: lessons from live cell imaging. *Biochim Biophys. Acta* **1799**, 15–27 (2010).
- Catez, F., Brown, D. T., Misteli, T. & Bustin, M. Competition between histone H1 and HMGN proteins for chromatin binding sites. *EMBO Rep.* **3**, 760–766 (2002).
- Rochman, M. et al. The interaction of NSBP1/HMGN5 with nucleosomes in euchromatin counteracts linker histone-mediated chromatin compaction and modulates transcription. *Mol. Cell* **35**, 642–656 (2009).
- Murphy, K. J. et al. HMGN1 and 2 remodel core and linker histone tail domains within chromatin. *Nucleic Acids Res* **45**, 9917–9930 (2017).
- Kato, H. et al. Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc. Natl Acad. Sci. USA* **108**, 12283–12288 (2011).
- Kalashnikova, A. A., Porter-Goff, M. E., Muthurajan, U. M., Luger, K. & Hansen, J. C. The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J. R. Soc. Interface* **10**, 20121022 (2013).
- Deng, T. et al. Functional compensation among HMGN variants modulates the DNase I hypersensitive sites at enhancers. *Genome Res* **25**, 1295–1308 (2015).
- He, B. et al. Binding of HMGN proteins to cell specific enhancers stabilizes cell identity. *Nat. Commun.* **9**, 5240 (2018).
- Nanduri, R., Furusawa, T. & Bustin, M. Biological Functions of HMGN Chromosomal Proteins. *Int. J. Mol. Sci.* **21**, 20449 (2020).
- Garza-Manero, S. et al. Maintenance of active chromatin states by HMGN2 is required for stem cell identity in a pluripotent stem cell model. *Epigenetics Chromatin.* **12**, 73 (2019).

25. Lane, A. A. et al. Triplication of a 21q22 region contributes to B cell transformation through HMGN1 overexpression and loss of histone H3 Lys27 trimethylation. *Nat. Genet.* **46**, 618–623 (2014).
26. Cabal-Hierro, L. et al. Chromatin accessibility promotes hematopoietic and leukemia stem cell activity. *Nat. Commun.* **11**, 1406 (2020).
27. Bustin, M. Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol. Cell Biol.* **19**, 5237–5246 (1999).
28. Heintzman, N. D. et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
29. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).
30. Crippa, M. P., Alfonso, P. J. & Bustin, M. Nucleosome core binding region of chromosomal protein HMGM-17 acts as an independent functional domain. *J. Mol. Biol.* **228**, 442–449 (1992).
31. Mardian, J. K., Paton, A. E., Bunick, G. J. & Olins, D. E. Nucleosome cores have two specific binding sites for nonhistone chromosomal proteins HMG 14 and HMG 17. *Science* **209**, 1534–1536 (1980).
32. Voigt, P. et al. Asymmetrically modified nucleosomes. *Cell* **151**, 181–193 (2012).
33. Trieschmann, L., Martin, B. & Bustin, M. The chromatin unfolding domain of chromosomal protein HMGM-14 targets the N-terminal tail of histone H3 in nucleosomes. *Proc. Natl Acad. Sci. USA* **95**, 5468–5473 (1998).
34. Prymakowska-Bosak, M. et al. Mitotic phosphorylation prevents the binding of HMGN proteins to chromatin. *Mol. Cell Biol.* **21**, 5169–5178 (2001).
35. Kim, J. M. et al. Linker histone H1.2 establishes chromatin compaction and gene silencing through recognition of H3K27me3. *Sci. Rep.* **5**, 16714 (2015).
36. Catez, F., Ueda, T. & Bustin, M. Determinants of histone H1 mobility and chromatin binding in living cells. *Nat. Struct. Mol. Biol.* **13**, 305–310 (2006).
37. Teif, V. B. et al. Linker histone epitopes are hidden by in situ higher-order chromatin structure. *Epigenetics Chromatin* **13**, 26 (2020).
38. Cao, K. et al. High-resolution mapping of h1 linker histone variants in embryonic stem cells. *PLoS Genet.* **9**, e1003417 (2013).
39. Karmodiya, K., Krebs, A. R., Oulad-Abdelghani, M., Kimura, H. & Tora, L. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* **13**, 424 (2012).
40. Kugler, J. E. et al. High mobility group N proteins modulate the fidelity of the cellular transcriptional profile in a tissue- and variant-specific manner. *J. Biol. Chem.* **288**, 16690–16703 (2013).
41. Koldin, B., Suckow, M., Seydel, A., von Wilcken-Bergmann, B. & Muller-Hill, B. A comparison of the different DNA binding specificities of the bZip proteins C/EBP and GCN4. *Nucleic Acids Res.* **23**, 4162–4169 (1995).
42. Di Giannardino, D. C. et al. KLF4 is involved in the organization and regulation of pluripotency-associated three-dimensional enhancer networks. *Nat. Cell Biol.* **21**, 1179–1190 (2019).
43. Alfonso, P. J., Crippa, M. P., Hayes, J. J. & Bustin, M. The footprint of chromosomal proteins HMGM-14 and HMGM-17 on chromatin subunits. *J. Mol. Biol.* **236**, 189–198 (1994).
44. Zhou, K., Gaullier, G. & Luger, K. Nucleosome structure and dynamics are coming of age. *Nat. Struct. Mol. Biol.* **26**, 3–13 (2019).
45. Shaytan, A. K. et al. Coupling between Histone Conformations and DNA Geometry in Nucleosomes on a Microsecond Timescale: Atomistic Insights into Nucleosome Functions. *J. Mol. Biol.* **428**, 221–237 (2016).
46. Kim, J., Lee, J. & Lee, T. H. Lysine Acetylation Facilitates Spontaneous DNA Dynamics in the Nucleosome. *J. Phys. Chem. B* **119**, 15001–15005 (2015).
47. Zhang, S. et al. Epigenetic regulation of REX1 expression and chromatin binding specificity by HMGNs. *Nucleic Acids Res.* **47**, 4449–4461 (2019).
48. Deng, T. et al. Interplay between H1 and HMGN epigenetically regulates OLIG1&2 expression and oligodendrocyte differentiation. *Nucleic Acids Res.* **45**, 3031–3045 (2017).
49. Yuan, W. et al. Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. *Science* **337**, 971–975 (2012).
50. Michael, A. K. et al. Mechanisms of OCT4-SOX2 motif readout on nucleosomes. *Science* **368**, 1460–1465 (2020).
51. Dodonova, S. O., Zhu, F., Dienemann, C., Taipale, J. & Cramer, P. Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function. *Nature* **580**, 669–672 (2020).
52. Paton, A. E., Wilkinson-Singley, E. & Olins, D. E. Nonhistone nuclear high mobility group proteins 14 and 17 stabilize nucleosome core particles. *J. Biol. Chem.* **258**, 13221–13229 (1983).
53. Chronis, C. et al. Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell* **168**, 442–459 e420 (2017).
54. Mowery, C. T. et al. Trisomy of a Down Syndrome Critical Region Globally Amplifies Transcription via HMGN1 Overexpression. *Cell Rep.* **25**, 1898–1911 e1895 (2018).
55. Deng, T. et al. HMGN1 modulates nucleosome occupancy and DNase I hypersensitivity at the CpG island promoters of embryonic stem cells. *Mol. Cell Biol.* **33**, 3377–3389 (2013).
56. Zhang, S. et al. HMGN proteins modulate chromatin regulatory sites and gene expression during activation of naive B cells. *Nucleic Acids Res.* **44**, 7144–7158 (2016).
57. Ausio, J., Dong, F. & van Holde, K. E. Use of selectively trypsinized nucleosome core particles to analyze the role of the histone “tails” in the stabilization of the nucleosome. *J. Mol. Biol.* **206**, 451–463 (1989).
58. Postnikov, Y. V. & Bustin, M. Analysis of HMGM-14/-17-containing chromatin. *Methods Mol. Biol.* **119**, 303–310 (1999).
59. Postnikov, Y. V. & Bustin, M. Reconstitution of high mobility group 14/17 proteins into nucleosomes and chromatin. *Methods Enzymol.* **304**, 133–155 (1999).

Acknowledgements

This project was funded by the Center for Cancer Research, Intramural Research Program of the National Cancer Institute, NIH and by Contract No. HHSN26120080001E from the National Cancer Institute. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. We thank Dr. Ravikanth Nanduri and Dr. Bing He for comments on this manuscript.

Author contributions

S.Z. and M.B. conceived and designed the study. S.Z., Y.P., T.F., and T.D. performed experiments and analyzed data. A.L. and S.Z. performed bioinformatic analyses. S.Z. and M.B. wrote the manuscript, M.B. coordinated the project.

Funding

Open Access funding provided by the National Institutes of Health (NIH).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-022-03099-0>.

Correspondence and requests for materials should be addressed to Michael Bustin.

Peer review information *Communications Biology* thanks Chitvan Mittal and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Patrick Murphy and Christina Karlsson Rosenthal.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2022