Histone H4 lysine 20 mono-methylation directly facilitates chromatin openness and promotes transcription of housekeeping genes

Muhammad Shoaib, Qinming Chen, Xiangyan Shi, Nidhi Nair, Chinmayi Prasanna, Renliang Yang, David Walter, Klaus S. Frederiksen, Hjorleifur Einarsson, J. Peter Svensson, Chuan Fa Liu, Karl Ekwall, Mads Lerdrup, Lars Nordenskiöld, & Claus S. Sørensen

Histone lysine methylations have primarily been linked to selective recruitment of reader or effector proteins that subsequently modify chromatin regions and mediate genome functions. Here, we describe a divergent role for histone H4 lysine 20 mono-methylation (H4K20me1) and demonstrate that it directly facilitates chromatin openness and accessibility by disrupting chromatin folding. Thus, accumulation of H4K20me1 demarcates highly accessible chromatin at genes, and this is maintained throughout the cell cycle. In vitro, H4K20me1-containing nucleosomal arrays with nucleosome repeat lengths (NRL) of 187 and 197 are less compact than unmethylated (H4K20me0) or trimethylated (H4K20me3) arrays. Concordantly, and in contrast to trimethylated and unmethylated tails, solid-state NMR data shows that H4K20 mono-methylation changes the H4 conformational state and leads to more dynamic histone H4-tails. Notably, the increased chromatin accessibility mediated by H4K20me1 facilitates gene expression, particularly of housekeeping genes. Altogether, we show how the methylation state of a single histone H4 residue operates as a focal point in chromatin structure control. While H4K20me1 directly promotes chromatin openness at highly transcribed genes, it also serves as a stepping-stone for H4K20me3-dependent chromatin compaction.
The landscape of chromatin accessibility is a continuum of different chromatin compaction states ranging from facultative and constitutive heterochromatin to more open, accessible euchromatin. The degree of chromatin compaction at any given genomic locus is determined by the nucleosomal packing as well as chromatin binding proteins. Nucleosomes—the fundamental repeating structural units of chromatin, consist of ~147 bp of DNA wrapped around an 8-mer of core histone proteins H2A, H2B, H3, and H4. Nucleosomal packing is influenced by histone composition and histone post-translational modifications at their N-terminal and C-terminal tails, which then modulates chromatin compaction states.

The histone H4 lysine 20 (H4K20) methylation pathway has previously been linked to the maintenance of a compact chromatin state. This pathway is highly cell cycle regulated as H4K20me1, catalyzed by the essential histone methyltransferase SET8, peaks in G2 and M phase. H4K20me1 is then further modified to higher methylation states (H4K20me2/3) by SUV4-20H1/2 enzymes in the G1 and S phase of the cell cycle. Remarkably, although H4K20me1 and H4K20me3 marks have both been linked to chromatin compaction and transcriptional repression, H4K20me1 is predominantly enriched within actively transcribing gene bodies. These regions are characterized by a more open chromatin folding state, hence, the H4K20me1 enrichment at these regions appears contradictory to its previously suggested role in chromatin compaction. Moreover, the potential effect of this histone modification on the biophysical properties and compaction of well-defined in vitro reconstituted chromatin fibers and its causal relationship to chromatin accessibility is currently not known. The histone H4-tail basic stretch mediates chromatin folding by interacting with an acidic patch of AAs of histone H4, enabling the interaction with an acidic patch of AAs of histone H4.

Histone H4 lysine 20 mono-methylation promotes chromatin accessibility at gene bodies. We first set out to explore how H4K20me1 and H4K20me3 methylation states vary with chromatin accessibility and compaction of different annotated regions of the genome over the cell cycle. To this end, we synchronized U2OS cells with a double thymidine block and harvested at the G1/S transition (t0), late S/G2 phase (t8), or the next G1 phase (t16) (Supplementary Fig. 1a–f). To measure genome-wide chromatin accessibility at these different phases of cell cycle, we performed ATAC-seq (Assay for Transposase-Accessible Chromatin with high-throughput sequencing). ATAC-seq relies on hyperactive transposase Tn5, which fragments chromatin genome-wide at accessible locations, followed by next-generation sequencing. Although ATAC-seq mainly captures nucleosome free regions, it has also been shown to enable identification of longer multi-nucleosomal reads and can be used to study a larger range of chromatin accessibility states at different types of genomic loci. To define H4K20me1 and H4K20me3 enriched loci we performed genome-wide profiling by chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) (Supplementary Fig. 1g). When comparing ATAC-seq and ChIP-seq levels at annotated genes, we found that genes with higher H4K20me1 levels had higher chromatin accessibility than genes with lower H4K20me1 levels (Fig. 1a, b), consistent with enrichment of H4K20me1 at actively transcribing genes. In contrast, H4K20me3 enriched loci generally displayed reduced chromatin accessibility (Fig. 1a, b). Notably, these accessibility profiles of H4K20me1 and H4K20me3 loci were preserved throughout the different cell cycle phases and in an asynchronous population of cells (Fig. 1a, b). Supplementary Fig. 1h shows the SET8-kd and H4K20me1 levels in asynchronous U2OS cells. As expected, H4K20me1 was enriched within intragenic and gene overlapping regions compared to intergenic loci, and globally it was negatively correlated to H3K9me3, a histone mark abundantly present in transcriptionally repressed loci. Furthermore, we observed a positive genome-wide correlation between H3K9me3 and H4K20me3 (Supplementary Fig. 1i). In conclusion, these observations indicate a marked association between H4K20me1 and chromatin accessibility in the gene bodies that is preserved throughout the cell cycle.

Additionally, we found a weakly positive correlation between H4K20me1 and accessibility at enhancers and transcription start sites (TSSes) across the different cell cycle phases (Supplementary Fig. 2a, c) and in asynchronously growing cells (Supplementary Fig. 2b, d). The relatively higher correlation at TSSes compared to enhancers indicates that H4K20me1 is more closely linked to genic transcription. Similar to our earlier findings in fission yeast, H4K20me1 was considerably more enriched at shorter genes than the H4K20me3 mark (Supplementary Fig. 3a–c). While both short and accessible genes were more enriched in H4K20me1 compared to larger or less accessible genes, H4K20me1 enrichment was generally more associated to accessibility than gene size (Fig. 1c and Supplementary Fig. 3d–f).

We next asked if H4K20me1 is functionally important for maintaining chromatin accessibility. To this end, we examined the accessibility profile of intragenic versus intergenic loci following SET8 depletion, which strongly reduced H4K20me1 levels (Supplementary Fig. 1d, f, h). In addition, we observed general changes in accessibility within gene bodies, where the less accessible 3' end had consistently reduced accessibility in contrast to the highly accessible 5' ends (Supplementary Fig. 4). Combined visualization of the changes in accessibility and in H4K20me1 mark revealed that intragenic loci with reduced accessibility after SET8 depletion were generally enriched in H4K20me1, implying that the accessibility of H4K20me1 enriched loci depends on SET8. This was evident at all phases of the cell cycle (Fig. 2a, left panels, bottom-right quadrant). These data point towards a dynamic role of SET8 in promoting chromatin accessibility of
intragenic loci throughout the cell cycle by establishing and maintaining H4K20me1 mark. In contrast, H4K20me1 levels were much lower at intergenic loci, and thus SET8 dependency was largely not observed in the intergenic loci (Fig. 2a right panels, bottom-right quadrant). Asynchronously growing cells displayed a similar SET8-dependent chromatin accessibility at H4K20me1 enriched loci (Fig. 2b), supporting the observation that SET8-dependent H4K20me1 primarily mediates chromatin accessibility at intragenic loci. We observed that in synchronized U2OS cells grown asynchronously, a clear positive correlation between H4K20me1 enrichment and chromatin accessibility at intragenic loci was dependent upon SET8 depletion (Supplementary Fig. 6a). Furthermore, we observed an overall weakly positive correlation between gene expression and chromatin accessibility at both intragenic and intergenic loci.

SET8-H4K20me1 driven chromatin accessibility promotes transcription of housekeeping genes. Our data indicated a role for H4K20me1 in maintaining a more accessible chromatin state, which was particularly evident in transcriptionally active regions of the genome. Thus, we investigated the correlation of H4K20me1 and H4K20me3 enrichment to gene expression. We observed a consistent positive correlation of H4K20me1 with the level of gene expression, whereas the inverse correlation was observed for H4K20me3 (Fig. 3a). However, this correlation was much weaker between the intergenic loci and the nearest gene (Supplementary Fig. 6a). Furthermore, we observed an overall weakly positive correlation between gene expression and chromatin accessibility at both intragenic and intergenic loci.

Fig. 1 Chromatin accessibility at genes bodies correlates with histone H4 lysine 20 mono-methylation throughout the cell cycle. a 2D-histograms showing the relationship between enrichment of selected H4K20 methylation states (Y-axis) and chromatin accessibility probed using ATAC-seq (X-axis) at human genes. X-axis shows log2 normalized H4K20me1 (left side) or H4K20me3 (right side) ChIP-seq levels relative to the levels of H4 control ChIP-seq, while Y-axis shows log10-transformed average FPKM-normalized ATAC-seq signal from synchronized U2OS cells at different cell cycle phases, i.e., G1/S (0 h—top, three replicates), S-G2 (8 h—middle, three replicates) and next G1 phase (16 h—bottom, four replicates). The color scales reflect the number of genes having a certain combination of H4K20me enrichment and ATAC-seq r-values show the Pearson’s correlation coefficients for the data in each plot. b 2D-histograms showing the relationship between enrichment of selected H4K20 methylation states (X-axis) and chromatin accessibility probed using ATAC-seq (Y-axis) at human genes as in a. Y-axis shows log10-transformed average FPKM-normalized ATAC-seq signal from four replicates of U2OS cells having a certain combination of ATAC-seq signal and size controlling opacity as indicated in the right side color scale. X-axis show log10-transformed average FPKM-normalized ATAC-seq signal from three replicates of synchronized U2OS cells 0 h after release.

### Notes
- The correlation coefficients are shown in the color scales of the histograms.
- The figure demonstrates the relationship between chromatin accessibility and histone H4 lysine 20 mono-methylation across different cell cycle phases.
- The data suggests a role for SET8 in maintaining accessible chromatin states at intragenic loci.
Based on our findings in Fig. 2, we then predicted that gene expression of H4K20me1 enriched genes is SET8 dependent and linked with the ability of the H4K20me1 mark to facilitate an open and accessible chromatin state. Indeed, after SET8-kd, the expression of H4K20me1-positive genes was significantly reduced (Fig. 3b, left panel). Notably, the loci that displayed loss of expression after SET8-kd and showed the most pronounced loss in accessibility had the highest average H4K20me1 enrichment (Fig. 3c, bottom-left quadrant). As expected, there was no observable relationship between H4K20me1 enrichment, chromatin accessibility changes at intergenic loci, and the expression changes of the nearest gene (Supplementary Fig. 6c, left panel and Supplementary Fig. 6d). In contrast to H4K20me1, H4K20me3 enriched loci did not show this relationship between gene expression and chromatin accessibility (Fig. 3b, right panel and Supplementary Fig. 6c, right panel). From these findings, we conclude that SET8-dependent H4K20me1 enrichment at active genes promotes an accessible chromatin state, which in turn is permissive for gene expression. Furthermore, H4K20me1-dependent chromatin accessibility at...
**Fig. 3** The SET8-H4K20me1 pathway driven chromatin accessibility promotes transcription. **a** 2D-histograms showing the relationship between enrichment of selected H4K20 methylation states (X-axes) and expression assayed by microarrays (Y-axis) at human genes. X-axes shows log2 normalized H4K20me1 (left side) or H4K20me3 (right side) ChIP-seq levels relative to the levels of H4 control ChIP-seq, while Y-axes shows log2 transformed expression levels from synchronized U2OS cells 16 h (G1 phase) after release. The color scales reflect the number of windows having a certain combination of K20me enrichment and expression. r-values shows the Pearson’s correlation coefficients for the data in each plot. **b** Colored volcano plots showing the three-way relationship between enrichment of selected H4K20 methylation states (color), gene expression difference (X-axis), and significance of the gene expression differences (Y-axis). Coloring shows log2 normalized H4K20me1 (left side) or H4K20me3 (right side) ChIP-seq levels relative to the levels of H4 control. The opacity in the diagram is controlled by the number of genes having the combination of X-axis and Y-axis values as indicated in the right-side color scale. X-axes show log2-transformed difference between SET8-kd over control-kd, and Y-axes show -log10 adjusted p-values from the analysis of the differential expression from two replicates of SET8-kd and three replicates of control-kd in synchronized U2OS cells analysed at 16 h (G1 phase) after release. **c** Colored 2D-histograms showing the genome-wide three-way relationship between enrichment of H4K20me1 (color), expression changes (X-axis), and ATAC-seq signal changes in SET8-kd over control-kd in synchronized U2OS. ChIP-seq and ATAC-seq signal was measured in 10 kbp windows that overlapped fully or partially with an annotated gene body. ATAC-seq signal and expression differences are plotted the log2 difference between SET8-kd over control-kd. Coloring shows the average log2 normalized H4K20me1 ChIP-seq levels relative to the levels of H4 control ChIP-seq as indicated in the right-side color scale. Opacity reflects the number of windows with a certain combination of expression change (Y-axis) and ATAC-seq change (X-axis) controlling opacity as indicated in the right-side color scale. Plots shows FPKM-normalized ATAC-seq signal from separate rounds of experiments as follows: synchronized U2OS cells at G1/S (0 h—top, three replicates), S-G2 (8 h—bottom, three replicates) and next G1 phase (16 h—bottom, four replicates), or U2OS cells grown asynchronously (rightmost, four replicates).

To this end, we clustered the intragenic H4K20me1 ChIP-seq signal according to the level and distribution along gene bodies in SET8-kd and control-kd conditions. Among the resulting clusters, the most significant transcriptional changes were observed at clusters 7–10 (Fig. 4a and Supplementary Fig. 6e, f), which contained generally small-sized genes (Supplementary Fig. 6g). The gene ontology analysis of these clusters revealed a strong enrichment of pathways involved in basal cellular functions, which mainly contained housekeeping genes (HK) (Supplementary Fig. 7). This subset of genes are constitutively expressed throughout the cell cycle and are critically required for cellular survival. Notably, a set of published HK genes were highly enriched in H4K20me1 mark as compared to non-HK genes (Fig. 4b). We then analyzed the three-way relationship between gene expression levels, chromatin accessibility and H4K20me1 enrichment at the HK genes and compared it to non-HK genes. Interestingly, we found that chromatin...
accessibility and gene expression of HK genes were strongly dependent on SET8 (Fig. 4c, top panels, bottom left quadrant). Consistently, this relationship remained constant throughout the cell cycle and in asynchronously growing cells. However, the chromatin accessibility and gene expression profile of non-HK genes did not show as strong a dependence on the presence of SET8 and H4K20me1 mark (Fig. 4c, bottom panels). Overall, the analysis of chromatin accessibility of HK versus non-HK genes
indicate that SET8-dependent H4K20me1 mark maintains chromatin accessibility of HK genes during different phases of cell cycle and hence, is an important determinant of the constitutive expression of HK genes.

Since, we observed that both accessibility and transcriptional activity of HK-genes are dependent on the presence of H4K20me1 mark, we next explore the possibility that transcriptional activity is the main driving force for maintaining the increased accessibility of H4K20me1 enriched intragenic loci. We, therefore, treated siSET8 and siCtrl U2OS cells with the transcriptional inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) or mock control for 2 h. DRB acts as a global inhibitor of RNA polymerase II (RNA Pol II)-dependent transcription by preventing activating phosphorylations of the RNA Pol II C-terminal domain. Supplementary Fig. 8a, b shows global transcriptional inhibition after treating U2OS cells ± SET8-kd with DRB inhibitor for 2 h. 5-ethyluridine (5-EU) pulse labeling was used to measure the transcriptional activity. For the analysis of chromatin accessibility in the presence of DRB, we generated ATAC-seq libraries from four replicates of each of the four conditions (Mock ± SET8 and DRB ± SET8). Consistent with our previous results, we observed that the accessibility of intragenic loci was dependent upon the presence of H4K20me1 in mock treated cells as loss of SET8 led to loss of accessibility of these loci (Fig. 4d top left and Supplementary Fig. 9a). To assess if the transcriptional output of genes impact their accessibility profiles, we next compared ATAC-seq profiles of siCtrl cells treated with either mock or DRB inhibitor. The results indicated that while decreasing the transcriptional output had some impact on the accessibility of genic loci, majority of the loci remained highly accessible (Fig. 4e and Supplementary Fig. 9b) suggesting that active transcription has quite moderate impact on the accessibility of genic regions. We then compared the accessibility profiles of siCtrl vs. siSET8 treated cells in the presence of DRB inhibitor. We reasoned that if transcriptional activity is the main mechanism responsible for maintaining the increased accessibility at H4K20me1 loci, then a combined treatment with DRB and siSET8 in U2OS cells would not be much different from DRB treatment alone. Notably, we observed an aggravated response in DRB-treated siSET8 cells as intragenic loci displayed a further decrease in accessibility in comparison to DRB treatment or siSET8 conditions alone. These data clearly suggest that the accessibility of intragenic loci is dependent upon the SET8-H4K20me1 pathway in a manner, that is not tightly linked to their transcriptional status (Fig. 2f, g and Supplementary Fig. 9c, d).

Histone H4 lysine 20 mono-methylation results in unfolding of 15-197-601 and 16-187-601 nucleosome arrays. The H4K20me1 mark is not known to be connected to any reader proteins that could immediately explain the increased chromatin accessibility or a decompacted chromatin state. Thus, we hypothesized that H4K20me1 may directly mediate chromatin decompaction, in particular, relative to the repressive H4K20me3 mark. To investigate this, we used in vitro reconstituted nucleosome arrays and performed analytical ultracentrifugation sedimentation velocity (AUC-SV) measurements. The sedimentation coefficient ($s_{20,w}$) measured in this assay is critically dependent on the chromatin fiber size and shape. This method therefore enables the detection of differences in the degree of compaction of nucleosome array constructs having varying H4K20 methylation states, and was previously used to establish the abrogating effect of acetylation at H4K16 on nucleosome arrays, thus demonstrating a mechanism for global decondensation of chromatin regions. Here, we used the AUC-SV method to elucidate whether the hypothesized effect of H4K20me1 on chromatin folding results in a change of the ability of the nucleosome array to reach the maximally compact state. We reconstituted chromatin fibers using a DNA array template comprising 15 repeats of 147 bp of the so-called Widom “601” high affinity nucleosome positioning sequence having a nucleosome repeat length (NRL) of 197 bp (15-197-601 DNA) as well as with NRL of 187 bp (16-187-601), which corresponds to linker DNA lengths (nucleosome–nucleosome spacing) of 50 and 40 bp, with recombinant Homo sapiens core histones (H2A, H2B, H3 and H4) (Supplementary Fig. 10a). We conducted the installation of methyl-lysine analogs according to established protocols (see Online Methods) and corroborated the successful installation of the mono-methylation, and tri-methylation constructs by mass-spectrometry (Supplementary Fig. 11a–c). We then applied the AUC-SV method to examine the folding of individual chromatin fibers in addition, constructs with NRL of 202 bp (12-202-601) were prepared and characterized in the same way.

We analyzed the AUC-SV experiments by the van Holde-Weischet method (Fig. 5a, d and Supplementary Fig. 10b–d). Under buffer conditions and in the absence of added Mg$^{2+}$ both methylated (H4K20me1 and H4K20me3) as well as unmodified...
(H4K20me0) arrays exhibited the behavior expected for an extended “beads-on-a-string” array conformation with a sedimentation coefficient ($s_{20,w}$) ~ 41-42S for 197 as well as for 187 arrays (Fig. 5b, e). Upon addition of Mg$^{2+}$, the H4K20me0 and H4K20me3 the 197 arrays folded gradually and finally reached the maximal degree of compaction at 1.4 mM Mg$^{2+}$ with a characteristic $s_{20,w}$ ~58S. Remarkably, H4K20me1 exhibited an abrogating effect on chromatin folding; the most compacted state displaying an $s_{20,w}$ value of ~53S at 1.4 mM Mg$^{2+}$ (Fig. 5b, c).

This effect of the H4K20me1 mark was somewhat more pronounced for Na$^+$-induced chromatin compaction, which displayed a maximal $s_{20,w}$ value of ~50S (Supplementary Fig. 10c). The results for the 187 NRL (Fig.5e, f) that corresponds to linker DNA length of 40 bp, mirror the findings seen for the 197 construct (with linker DNA length 50 bp), and confirms that this effect is conserved for this shorter NRL that also has a 10n bp nucleosome spacing. These observations suggest that the mono-methylation resulted in a marked disturbance of the H4-tail-mediated nucleosome–nucleosome interaction, which disrupted stacking of folded arrays. The reason that tri-methylation did not
show such an effect was likely due to the introduction of the highly hydrophobic group of the K20 side chain. This enables strong hydrophobic interactions with the surface of the neighboring nucleosome, which therefore can maintain the H4 tail-mediated nucleosome–nucleosome stacking characteristic of folded arrays.

We additionally investigated 12–202–601 nucleosome arrays that have a 10n+5 bp rotational setting (linker DNA length 55 bp). These arrays display considerably less pronounced folding in the presence of Mg2+ and there is no effect of the H4K20 methylations (see Supplementary Fig. 12). The result is not unexpected since the changed rotational setting would force the nucleosome–nucleosome stacking orientation to be shifted by 180 degrees. This disrupts the H4 tail interaction with the acidic patch of the neighboring nucleosome leading to less compact arrays, which likely lack the H4 tail mediated interaction that is disrupted by H4K20me1 methylation in the case of 187 and 197 arrays. A study of the Mg2+ induced compaction behavior previously showed such a reduced folding for 172 compared to 177 nucleosome arrays43. It was also recently shown that 601 chromatin fibers close to 10n+5 bp rotational setting displayed reduced stacking energy and fibers that are more flexible, compared to constructs close to a 10n bp rotational setting40.

The observation that the H4K20me3 modification had negligible effect on fiber compaction was different from an earlier report41 for 12–207–601 nucleosome arrays, where H4K20me3 resulted in somewhat more compact and oligomerized nucleosome arrays compared to the disruption in folding seen for H4K20me1 in the present work. This difference may be a result of different sample preparation (the previous study used Xenopus Laevis histones) or due to the longer nucleosome repeat lengths applied in the experiments, which has previously been shown to result in less compact nucleosome arrays in the presence of Mg2+42. The choice of a NRL of 187 and 197 in the present work was made based on the analysis by Widom of a large number of measurements of NRL in eukaryotic cells, which deduced that values close 187 and 197 are the most abundant36, although it should be noted that a variation in observed NRL in the range between 190–200 bp is common in human cells42.

Next, we conducted precipitation assays (PA) to examine the efficiency of Mg2+ cations to induce self-association of arrays due to fiber–fiber interactions34,35. As we established from the AUC-SV measurements, addition of salt to a buffer solution of arrays resulted in a gradual folding of individual arrays until the fibers reached the maximally folded state described above. Further addition of salt to maximally folded arrays resulted in array self-association (oligomerization) (Supplementary Fig. 13a–c). We observed a modest but significant difference whereby the H4K20me1 arrays required the highest concentrations of Mg2+ necessary to result in fiber–fiber interaction and self-association (Supplementary Fig. 13a–c). This result was consistent with the AUC-SV data and suggested that mono-methylation on H4K20 disrupts the attraction between fibers caused by histone tail–DNA interactions43. Taken together the AUC-SV and PA data strongly suggest a direct role of H4K20me1 mark in preventing the maximal folding of nucleosome arrays, as well as reducing fiber–fiber interactions thereby creating a more open chromatin state. Importantly, this in vitro analysis supports our cell biological data where the H4K20me1 mark functionally correlated with highly accessible chromatin loci (Fig. 1). It should be noted that recent investigations suggest that chromatin fiber self-association (oligomerization) may be the main contributor to chromatin structure in vivo, which may lack distinctly folded 30 nm fibers44. The present data illustrates how the H4K20me1 mark affects both the folding of individual arrays as well as fiber self-association.

Solid-state NMR reveals change in conformational state of histone H4 tail upon K20 mono-methylation. The AUC-SV results showed that mono-methylation of histone H4K20 results in array unfolding. We therefore reasoned that this could be due to an effect on the interaction of H4 tail with neighboring nucleosomes, which mediates nucleosome–nucleosome stacking and facilitates fiber compaction. Accordingly, we hypothesized that H4K20me1 results in changes in the dynamic conformational state of the histone H4 tail. To investigate this at the molecular level, we performed solid-state NMR (ssNMR) studies of the histone H4 in folded and precipitated 197 nucleosome arrays and compared the dynamic behavior of H4K20me0, H4K20me1 and H4K20me3 arrays. The ssNMR experiments enable AA site-specific information on the dynamic and structural properties of the histone to be assessed and compared for arrays with different H4K20 methylation states.

First, in order to probe the conformation and dynamics of the H4 tail in the tightly compacted and precipitated arrays, we collected 1H,13C/15N correlation ssNMR spectra in order to detect those AAAs in the H4 tail region, which exhibits significant mobility compared with the rigid core region45,46. By using the known averaged chemical shift values of AAs combined with three-dimensional (3D) solution-state NMR experiments of the histone H4 “601” nucleosome core particle47, we established the tail peak chemical shift assignments. Interestingly, two peaks, 3.50–62.4 ppm and 3.45–62.8 ppm, were observed in the 1H–13C spectra of the nucleosome arrays containing H4K20me0 or H4K20me3 (Fig. 6a–d). These peaks could be uniquely assigned to the AA Valine 21 (V21) 1Hα–13Cα because this is the only Valine in the H4 N-terminal tail (Fig. 6a). Furthermore, we unambiguously assigned the H4V21 peak, which made it unique in contrast to the other AA in the tail region that all displayed considerable overlap. Additionally, V21 is located in the middle of the basic AA stretch 16–23 between the more flexible residues towards the N-terminal tail and the significantly more rigid ones closer towards the globular domain, which makes it exclusively positioned to report on the tail conformational dynamics. The two distinct V21 1Hα–13Cα peaks in the 15-mer nucleosome arrays containing H4K20me0 or H4K20me3 strongly suggested the presence of two conformations of the tail, referred to as State 1 and State 2, respectively. These two peaks potentially originate from two different tail positions that may correspond to the tails of the two H4 copies in a histone octamer (HO). This is consistent with crystal structure studies of the nucleosome core particle demonstrating that only one of the two H4 tails in the HO is involved in the interaction with the H2A-H2B acidic patch48. In contrast, as shown in Fig. 6b, d, the V21 1Hα–13Cα peak intensity of State 2 was negligible in the H4K20me1 containing array, indicating that the H4 tails predominantly existed as State 1, suggesting that the interaction between one H4 tail with the H2A-H2B acidic patch is disrupted and, therefore, both tails possess the same conformational state. This observation implies that the partial unfolding of the H4K20me1 array leads to a more open and dynamic chromatin fiber, a result that is corroborated by the more dynamic H4 core region (see below).

Secondly, we acquired dipolar-based ssNMR spectra of compact folded arrays to identify the H4 AA sites within the rigid core region in order to establish any dynamic changes introduced by the different H4K20 methylation states. First, we noted that the overall structures of the H4 globular domains in the three different nucleosomes arrays were identical as evidenced by the 2D spectra (Supplementary Fig. 14a–c). By extracting the peak intensities of the so-called NCA (N to alpha Ca) and NCO (N to carbonyl CO) spectra (Fig. 6e and Supplementary fig. 15a, b), we obtained information about the internal dynamics of the
H4 protein backbone at the μs-ms timescale. In comparison with the H4K20me0 nucleosome array, the globular domain of the H4K20me1 array showed more uniform peak intensities across the protein and reduced relative intensities for several residues primarily residing in the LN and L1 regions and in the adjacent stretch of the α1 and α2 regions (Fig. 6e and Supplementary Fig. 15a, b). This implied that these core regions of H4K20me1 were more mobile compared to the H4K20me0 arrays, likely related to a higher flexibility of DNA. In contrast, compared with the H4K20me0 arrays, the H4K20me3 construct

---

1. **H4** protein backbone at the μs-ms timescale. In comparison with the H4K20me0 nucleosome array, the globular domain of the H4K20me1 array showed more uniform peak intensities across the protein and reduced relative intensities for several residues primarily residing in the LN and L1 regions and in the adjacent stretch of the α1 and α2 regions (Fig. 6e and Supplementary Fig. 15a, b). This implied that these core regions of H4K20me1 were more mobile compared to the H4K20me0 arrays, likely related to a higher flexibility of DNA. In contrast, compared with the H4K20me0 arrays, the H4K20me3 construct.

---

2. **H4** protein backbone at the μs-ms timescale. In comparison with the H4K20me0 nucleosome array, the globular domain of the H4K20me1 array showed more uniform peak intensities across the protein and reduced relative intensities for several residues primarily residing in the LN and L1 regions and in the adjacent stretch of the α1 and α2 regions (Fig. 6e and Supplementary Fig. 15a, b). This implied that these core regions of H4K20me1 were more mobile compared to the H4K20me0 arrays, likely related to a higher flexibility of DNA. In contrast, compared with the H4K20me0 arrays, the H4K20me3 construct.
displayed enhanced relative peak intensities for many residues particularly in L2 and at the end of the C-terminus, which suggested a somewhat more rigid H4 globular domain, possibly correlated with less mobile DNA.

These above results are remarkable as they show that a single H4K20 mono-methylation, which globally causes unfolding of the chromatin fiber, at the molecular level is coupled with both a change in the conformational dynamics of the tail itself as well as it propagates change to the histone H4 globular core dynamics. Taken together, mono-methylation of H4K20 in the nucleosome array uniquely changes the H4 tail conformational population, shifting it from two states to one state as well as resulting in a more flexible histone. Trimethylation of H4K20 does not affect the H4 tail conformation, which remains similar to that of H4K20me0 arrays, albeit resulting in more rigid H4 core region.

Based upon the cell biological and structural data presented here, we present a model outlining how the SET8-H4K20me1 pathway impacts the structural and functional characteristics of HK genes (Fig. 7). We propose that HK genes are highly enriched in H4K20me1, which owing to its structurally more dynamic nature dictates a more open and accessible local chromatin environment. This type of chromatin structure in turn facilitates constitutive expression of HK genes. However, in the absence of SET8, HK genes lose H4K20me1 mark leading to a more static, compact, and transcriptionally less permissive chromatin environment. In conclusion, our findings highlight the structural and functional demarcation between the closely linked H4K20me states. We show that the cell cycle regulated H4K20me1 accumulation induces a more accessible, less compact chromatin that facilitates transcription. The mono-methylated state is a
stepping stone towards H4K20me3 in a subset of genomic regions, which in turn may induce a more compact chromatin folding state acting as an impediment to transcriptional activities. While histone methylation is a well-known recruitment mechanism for protein binding, our cell biological and biophysical data provide strong evidence that H4K20 mono-methylation can also alter the basic biophysical properties of chromatin and in turn has profound effects on crucial parts of the genome. Our research, therefore, opens up possibilities for further investigations into the role of additional histone methylation marks in directly contributing to chromatin accessibility regulation.

Methods

Cell culture and treatments. U2OS cells were obtained from ATCC while Mouse embryonic fibroblasts were a kind gift from Prof. Dr. Gunnar Schotta, Biomedical Center and Center for Integrated Protein Science, Ludwig-Maximilians-University, Munich, Germany. Both cell lines were cultured in Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum and 1% Penicillin. Cells were cultured for another 17 h. Cells were treated with released in fresh medium without thymidine for 10 h. Two millimolar thymidine was then added and cells were cultured for another 17 h. Cells were treated with either scrambled siRNA sequence (siCtrl) or siRNA against SET8 (siSET8) during the second block 6 h prior to second release (0–6 h). Cells were then washed two times with PBS and either harvested (t0–G1/S) or released in fresh medium. For analyzing cells in the subsequent S/G2 and next G1 phase, cells were collected after 8 (t0–late S/G2) and 16 h (next G1). For experiments in asynchronously cultured U2OS and MEFs, cells were plated and reverse transfected with either control or SET8 siRNA for 6 h. Fresh media was then added, and cells were incubated for up to 36–48 h (U2OS) or 72 h (MEF) in total from the time of transfection. siRNA transfections were performed with 20 nM siRNA duplexes using Lipofectamine® RNAiMAX (Invitrogen), according to the manufacturer’s instructions. The siRNA sequences used for knockdown are (5′-3′): SET8 (GUAGGGAGGCCACAAAGU) and SET8 mouse/human (GGUGACAGUGUGGAAGGUGAAG). U2OS cells were treated either with Mock (DMEM only) or with transcriptional inhibitor DRB (5′-6-Deoxy-5-fluoro-1,2,3-triazolo[4,5-b]pyridine-6-carboxamide) for 10 min at RT. Coverslips were washing 3 times with PBS and either harvested (t0–36 h) or released in fresh medium. For nascent RNA labeling, U2OS cells were pulse labeled with 5-ethyluridine (5-EU) (5 mM) for 1 h before fixation.

Flow cytometry. Cells were fixed in 70% ethanol and DNA was stained using 0.1 mg/ml propidium iodide containing RNase for 30 min at 37 °C. Flow cytometric analysis was performed on FACSCalibur using the CellQuest Pro software (BD). Data were analyzed using the FlowJo software (v10.6.2, TreeStar). At least 10,000 cells were analyzed for each sample. FL2-A was used to gate all positive cells. A narrow diagonal gate was used to select single population. All cells outside of this gate were selected out so they represent either doublets or dead cells. The cell cycle profiles shown in Supplemental Fig. 1c, e represents the cell population from this single cell gate.

Immunoblotting and antibodies. Cells were lysed on ice cold EBC-buffer (150 mM NaCl; 50 mM TRIS pH 7.4; 1 mM EDTA; 0.5% NP-40/leupeptin) containing protease inhibitors (EDTA free-Protease inhibitor cocktail, 1 mM PMSF, 50 mM NaF) actions occurred at RT. The lysates were sonicated using Bioruptor Plus sonication in cytosolic lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 5 mM MgCl2, 0.5 mM EGTA, 0.1% Triton X-100, 0.2% NaN3, 0.5 mM EDTA) for 2 min. 5U of Micrococcal nuclease (0.2 U/µl; Sigma Aldrich) was added and incubated at 37 °C for 3 min. The nuclei pellet was resuspended by adding PBS-T supplemented with secondary antibodies (1:2000), anti-H4K20me1 (Abcam; ab9051—5 µg/sample), anti-H4K20me3 (Abcam; ab9053—5 µg/sample), anti-H4 PAN (Abcam; ab9051—5 µg/sample). DNA was then purified using QIAquick spin column (QIAGEN). Sequencing was done on a NextSeq500 machine with read length of 76 bases. Following antibodies were used for ChIP: anti-H4K20me1 (Abcam; ab9051—5 µg/sample), anti-H4K20me3 (Abcam; ab9053—7 µg/sample), anti-H4 PAN (Millipore; 04–858—5 µg/sample).

Quantitative image-based cytometry (QIBC). Images used for QIBC were obtained automatically with the ScanR acquisition software controlling a motorized Nikon IX83 wide-field microscope. The system was operated with filter cubes compatible with DAPI and Cy5 fluorescent dyes, a Spectra X-LIGHT engine illumination system with six color LEDs and emission filters, and a Hamamatsu Camera Orca Flash 4.0 V2. An Olympus Universal Plan Super Apo 10 x Objective was used for all QIBC data. Images were processed using the ScanR image analysis software. BITCO SpotFire software was used to plot total nuclear pixel intensities for DAPI (Arbitrary units: arb. u.) and mean (total pixel intensities divided by nuclear area) nuclear intensities (arb. u.) for EU intensity in color-coded scatter diagrams in a flow-cytometry-like fashion.

ChIP-qPCR. Twenty million cells cultured U2OS cells were fixed with 1% formaldehyde on tissue culture plates at room temperature (RT). Fixation was stopped by adding of glycine to final a concentration of 0.125 M and incubate for 5 min. Plates were rinsed twice with 1 x PBS at RT. PBS was completely aspirated from the plate and cells were harvested in SDS Buffer (100 mM NaCl, 50 mM Tris-Cl, pH8.0, 5 mM EDTA, pH 8.0, 0.2% NaN3, and 0.5% SDS). Cells were pelleted by spinning for 10 min at 500 g. The cell pellet is resuspended in appropriate volume of ice-cold IP Buffer for sonication (IP buffer = 1 volume SDS Buffer 0.5 volume Trition Dition Buffer (100 mM Tris-Cl, pH 8.6, 100 mM NaCl, 5 mM EDTA, pH 8.0, 0.2% NaN3, 5% Triton X-100)). Samples were sonicated to an average length of 300–1000 bp using bath-based Bioruptor sonicator (10–12 cycles—high mode—30 s ON/30 s OFF) at room temperature (RT). After verifying the correct fragment size, chromatin was quantified using Bio-Rad Protein Assay Kit (Cat #5000002). For each individual ChIP, 0.5 mg chromatin was precleared using 40 µl of Protein A Sepharose beads (Protein A Sepharose® Fast Flow—GE Healthcare). One percentage input was taken out after pre-clearing and protein A Sepharose beads were incubated overnight in 5 µl of IP Buffer at 4 °C and rotation at 4 °C. BSA blocked Protein A Sepharose beads were added after ON incubation with primary antibody and samples were rotated for 3–4 h at 4 °C. Beads were subsequently collected by spinning at 1000 g for 2 min and washed three times with low salt wash buffer (150 mM NaCl, 20 mM Tris-HCl, pH 8.0, 5 mM EDTA, pH 8.0, 5% w/v sucrose, 0.2% NaN3, 1% Triton X-100, 0.2% SDS), high salt wash buffer (0.1% w/v) deoxycholic acid, 1 mM EDTA pH 8.0, 50 mM HEPES, pH 7.5, 500 mM NaCl, 1% (v/v) Triton X-100 and 0.2% NaN3, with LiCl wash buffer (0.5% w/v) deoxycholic acid (sodium salt), 1 mM EDTA pH 8.0, 250 mM LiCl, 0.5% (v/v) NP-40, 10 mM Tris-HCl pH 8.0 and 0.02% Nonidet P-40 (NP-40, in combination with TE buffer) was performed with TE buffer for ChIP-seq purposes. Samples and beads were then resuspended in decrosslinking buffer (1% SDS, 100 mM NaHCO3) and shaken at 65 °C for 10 min in thermoshaker (Eppendorf® Thermomixer Compact—T1317). DNA was then purified using QIAquick PCR purification kit (Cat No.79280). Chopped DNA was quantified using Qubit DNA HS Kit and 10 ng DNA was taken for prepping libraries for next generation sequencing using i8Qeal Chip-seq kit (Cat. # Q01005059), and sequencing was done on a Hiseq2000 machine with read length 55 bases. Native ChIP was performed for histone modifications in MEFs wherein 3 million MEFs were harvested by trypsinization. Following trypsinization, all steps were performed on ice unless otherwise. Cells were washed with PBS-T supplemented with 5% skimmed milk powder. Proteins were visualized on films using secondary horseradish peroxidase conjugated antibodies and antibodies against histone H4 (Gene CHEM), (0.5 mM) Sigma Aldrich) for 10 min at RT. Coverslips from outside of this gate were selected out so they represent either doublets or dead cells. The cell cycle profiles shown in Supplemental Fig. 1c, e represents the cell population from this single cell gate.

Immunostaining. Cells grown on 12 mm coverslips were washed twice with PBS, and then fixed with 4% formaldehyde for 10 min at RT. Cover slips were subsequently incubated in blocking solution (1 x PBS containing 0.1% Triton X-100 and 3% bovine serum albumin) for 1 h at RT. For EU detection Click-it reaction was performed by incubating the fixed cells in Click-it buffer (100 mM Tris-HCl pH 8, 2 mM CuSO4, 1 µg Alexa Fluor 647 Azide (Life Technologies), and 100 mM ascorbic acid for 30 min at RT. Coverslips were washed three times with 1 x PBS and incubated in the above solution. Sigma Aldrich) for 10 min at RT. Coverslips were washed three times with 1 x PBS and dipped in distilled water, placed on 3 mm paper to dry, and mounted on 5 µl fluorescent mounting media (Dako).
ATAC-seq. ATAC-seq was performed as originally described by Buenrostro et al.\(^{26,27}\) with a few modifications. Briefly, 50,000 cells were washed with 50 μl ice-cold 1× PBS followed by centrifugation at 500×

 cold lysis buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM MgCl₂, 0.1% NP40, 0.1% Tween-20 and 0.01% Digitonin). Cells were gently pipetted up and down and incubated for 3 min. Five hundred microliter of Wash buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM MgCl₂ and 0.1% Tween-20) was added, followed by centrifugation at 500×g for 10 min using a refrigerated centrifuge to collect nuclei. The nuclei pellet was resuspended in the transposase reaction mixture containing 25 μl of 2×TD buffer, 2.5 μl transposase (Illumina, FC-121-1031), 0.1% Tween-20 and 0.01% Digitonin and nuclease-free water up to 50 μl. The nuclei were resuspended in setting up an array. 100 kilobasespairs at 500 bp were obtained by importing the chromatin for each chromosome in the reference genome as a "Regionset" and using "Regionset/Modify"-tool and the settings "Homogenize and fragment regions to" and "10,000" bp. For genes and 10 kb windows, values were quantified using the "Quantify"-tool and default settings (resulting in values being normalized to FPKMs, Fragments Per Kilobase of transcripts and read (FPKM) read, except that start and end positions were set to region "Start" and "End", respectively, and offsets set to "0" bp. For TSSes and enhancers, values were quantified using the "Quantify"-tool and default settings (resulting in FPKM normalized values), except that start and end offsets were removed. As a result of the exclusion of reads and selecting the settings "Border to border of the regions" followed by gating out subsets with border distances =<0 or >0 for overlapping, and non-overlapping regions, respectively. A similar strategy was used to obtain windows of size 1kbp and "Window size" set to 200%. Meta-Gene plots of gene quartiles were generated using R (https://www.R-project.org)\(^{31}\) and simple calculations and database functions were performed in Microsoft Excel.

**Visualization of ATAC-seq, ChIP-seq, and expression data.** When nothing else is stated, visualization was done using EqaSeq (v. 1.1111)\(^{39}\) and integrated tools and exported for layouting using the "Export snapshot as pdf" tool in the Beta-testing panel. Color scales were developed by Smith and van der Walt as well as Harrow and Brewer\(^{40}\) (see https://bioconductor.org/). 2D-histograms were obtained using the "Scatter"-tool, and colored 2D-histograms were obtained using the "ZScatter"-tool with the number of bins on the axes set to 100 and "Count" adjusted as indicated in the "Plot Settings". Graphs of the average distribution of signal along gene bodies were obtained using the "Av. Track" tool, followed by selecting relevant graphs and exported for layouting using the "Overlay"-tool. Box plots were obtained using R (https://www.R-project.org)\(^{35}\), simple vertical 1D heatmaps of numerical values were generated using the "Par-Map"-tool, and Heatmaps of ChIP-seq signal distribution along several loci were generated using the "HeatMap" (for heatmaps of one ChIP-seq data). "Ratio-Map"-tool and redundant graphics along gene bodies were obtained using the "2D-histograms" tool. Violin plots were generated using the "Histogram"-tool. Pearson's correlation coefficients in 2D-histograms of NGS data generally resulted in p-values at the lowest possible number (p < 2.2e-16) that can be obtained using the "cor.test()" in R (https://www.R-project.org)\(^{35}\).
The pET-3a plasmids containing the histone sequences were transformed and overexpressed in E. coli BL21 (DE3) pLysS cells in 2 T media. Crude WT histones were purified by 26/60 Sephacryl S-200 column (GE Healthcare) followed by the Resource S cation-exchange column (GE Healthcare). The expression of H4K20C mutant was similar to that of WT histones. H4K20C histone purified by gel-filtration (Sephacryl S-200 column) was further purified by Vydac C4 preparative HPLC (Grace). To produce 15C-labeled, 14N-labeled histones, cells were grown in 2 T media until OD600 reached 0.5. The cells were pelleted down and resuspended in equal volume of M9 minimal media supplemented with 0.4% D-glucose and 0.1% 15N ammonium chloride (CEI, USA), micro nutrients and BME (Sigma-Aldrich). After 1 h, acetate was induced with 0.4 M IPTG at OD600 of 0.7 for overexpression for 3 h at 37 °C. The 15C-labeled, 14N-labeled WT H4 and H4K20C were purified following the same procedures as the unlabeled ones.

**Installation of methyl-lysine analogs.** To generate histone H4 with mono-methylation and tri-methylation on Lys20, we followed the modified published protocol, which results in the replacement of the lysine side chain with the methylated analog that has an S atom at the gamma position (denoted K20 methylation)37. Previous in vitro investigations established that the K20 methylation analogs are not compatible with histones. The methylated analogs were obtained from the respective distributors with respect to biochemical and biophysical properties37,41. For monomethyl histone installation, briefly, 10 mg of pure H4K20C protein was dissolved in 900 μl of 0.1 M alkaline buffer (6 M Gdn·HCl, 1 M HEPE, 100 mM D/L methionine, pH 7.8), and incubated in 37 °C for 1 h after the addition of 20 μl of freshly prepared dimethylated (D2MeH4) trimethylammonium chloride (96% purity, Sigma-Aldrich) to the solution. After 1 h, the solution was added to the sample. After 30 min, 250 μl of (2-bromoethyl)trimethylammonium bromide was added. The reaction took place at 60 °C. After 5 h, another 20 μl of 100 mM DTT and 2.6 mg of (2-chloroethyl)trimethylammonium chloride was added, and the reaction proceeded overnight at room temperature (Grace).

For trimethylation installation, briefly, 40 mg of H4K20C protein was dissolved in 4.9 ml of 0.1 M alkaline buffer. Hundred microliter of freshly prepared 1 M DTT was added to the solution. After 10 min, 300 mg of (2-bromomethyl)trimethylammonium bromide was added. The alkylation reaction took place at 50 °C. After 5 h, another 50 μl of 1 M DTT and 147 mg of (2-bromomethyl)trimethylammonium bromide was added. The reaction proceeded for another 3 h at 50 °C.

For both mono-methylation and tri-methylation, the reaction was quenched with Trifluoroacetic acid (TFA), after which the mixture was dialyzed against water (containing 0.1% TFA) overnight. The final product was obtained after lyophilization. The molecular weight of the final product was confirmed by MALDI-TOF MS.

**Histone octamer refolding, DNA preparation, and chromatin preparation.** Histones (hH2A3aH2B3aH3aH4) were produced in unifolding buffer (7 M Gdn·HCl, 10 mM Tris·HCl, pH 7.5) at molar ratio of 1:2.1:2.1:1 and dialyzed against refolding buffer (10 mM Tris·HCl, 2 M NaCl, 1 mM EDTA, 5 mM beta-mercaptoethanol, pH 7.5) overnight. The histone octamer (HO) was purified by gel filtration chromatography using HiLoad 16/60 Superdex 200 pg column (GE Healthcare), and 15-601-197 DNA was obtained by HiPrep 26/60 Sephacryl S-500 HR column (GE Healthcare). The purity of the fractions was checked by 0.7% agarose gel.

The 15-601-197 DNA (15 repeats of the Widom 601) with full affinity nucleosome positioning sequence with 197 bp NRL) originally constructed in published protocols20,23. 15-197 DNA, purified by 0.7% agarose gel. The purity of the fractions was checked by Vydac C4 preparative HPLC (Grace). The molecular weight of the final product was confirmed by MALDI-TOF MS.

**NMR sample preparation.** Reconstituted 15-197 arrays were concentrated with an Amicon concentrator device (50 kDa, Millipore) to 7–9 mg/ml and treated with 5 mM MgCl2 (final concentration). The precipitated samples were transferred into a rotor packing device (Gioffet Biotech) with a 19 mm SSNMR rotor. Ultracentrifugation at 100,000 g was performed for 5–8 h to bring down the nucleosome array aggregates into the 19 mm SSNMR rotor.

**Solid-state NMR experiments.** All ssNMR experiments were conducted with an 18.8T Bruker Advance III HD spectrometer equipped with a 1.9 mm HCN magic angle spinning (MAS) probe. The sample temperature was maintained at 11–13 °C, which was calibrated externally with ethylene-glycol42. The 13C chemical shift was referenced at the DSS scale using adamantane with the downfield signal set at 40.48 ppm. The 1H and 14N chemical shifts were indirectly calculated based on the 13C reference 65. All experiments were performed with a MAS frequency of 17857 Hz. The 90° pulse lengths of 1H, 13C, and 14N were 2.4 μs, 3.2 μs, and 4.15 μs, respectively. In the 2D J-coupled 1H–13C correlation experiments, the spectral width and acquisition time were 220.22 ppm and 19.7 ms in the direct dimension, and 3.72 ppm and 16.1 ms in the indirect dimension. The spectral width and acquisition time used in the 2D 13C–13C DARR experiments were 266.23 ppm and 15.5 ms in the direct dimension and 266.23 ppm and 8.4 ms in the indirect dimension. In the 2D NCA experiments, the spectral width and acquisition time were 266.25 ppm and 14.4 ms in the direct dimension, and 44.0 ppm and 12.3 ms in the indirect dimension. In the 2D NCO experiments, the spectral width and acquisition time used were 88.74 ppm and 14.4 ms in the direct dimension and 36.7 ppm and 12.1 ms in the indirect dimension. The NCA and NCO transfer was achieved by SPECIFIC-CP65. In all experiments, the recycle delay was set as 1.5 s. All ssNMR data was processed using nmrPipe65 and analyzed with Sparky (T. D. Goodard and D. G. Kneller, University of California, San Francisco).

**Data availability**

The data that support this study are available from the corresponding authors upon reasonable request. The ChIP-seq data generated for this publication have been deposited in NCBI’s Gene Expression Omnibus66 and are accessible through GEO Series accession number GSE156216. ATAC-seq data from asynchronous cells, ATAC-seq data from the 0 h (t0), and from the 8 h (t8) time points are accessible through GEO Series accession number GSE156215 and ATAC-seq data from DRB treated cells and controls are available through GEO Series accession number GSE173313. ATAC-seq data from the 16 h time point were previously published11 and are available at accessible through GEO Series accession number GSE188606. Previously published H3K9me3 and input data67 were downloaded from the NCBI GEO database69 from GEO Series accession number GSE31755. Rfseq gene annotations68 were acquired from the UCSC table browser89. Enhancers were obtained from the Fantom 5 consortium90, and a list of human housekeeping genes was obtained from31. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE91 partner repository with the dataset identifier PXD007150. Source data are provided with this paper.

Received: 26 August 2020; Accepted: 15 July 2021; Published online: 20 August 2021

**References**

1. Bowman, G. D. & Poirier, M. G. Post-translational modifications of histones that influence nucleosome dynamics. Chem. Rev. 115, 2274–2295 (2015).
2. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. Cell Res. 21, 381–395 (2011).
3. Hiragami-Hamada, K. et al. Dynamic and flexible H3K0me3 bridging via HP1beta dimerization establishes a plastic state of condensed chromatin. Nat. Commun. 7, 11310 (2016).
and the research committee at the Karolinska Hospital, ChIP samples were prepared by core facility at KI, which is supported by the board of research at the Karolinska Institutet. ChIP-sequencing was performed at the Bioinformatics and Expression Analysis (BEA) laboratory. All ssNMR experiments were performed at the Nanyang Technological University (NTU) Center of High-Field NMR Spectroscopy and Imaging. We also acknowledge the NTU Institute of Structural Biology (NISB) for supporting this research.

The Singapore Ministry of Education Academic Research Fund (AcRF) Tier 2 Research Foundation for grant DNRF115 funding the Center for Chromosome Stability. Benzon foundation, and the Danish Cancer Society. We thank the Danish National Cancer Association for financial support. Work in the CSS laboratory was supported by the Novo Nordisk Foundation, the Danish Cancer Society, and the European Research Council (ERC) Project 844955 Chromatin Catapult. M.S., Q.C., and X.S. designed and carried out experiments, wrote the manuscript, and conducted data analysis. C.P. contributed to the NMR experiments. R.Y., and C.F.L. contributed to the synthesis and analysis of modified histones. N.N., D.W., J.P.S., and H.E. contributed to the ChIP-seq and ATAC-seq assays. D.W., K.S.F., and H.E. contributed to the gene expression analysis. All authors proofread and contributed to the manuscript. M.L., L.N., and C.S.S. oversaw the study, conducted data analysis, wrote the manuscript, and helped design the experiments.

**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/s41467-021-25051-2.

**Correspondence**
and requests for materials should be addressed to M.L., L.N. or C.S.S.

**Peer review information**
Nature Communications thanks the anonymous reviewers for their contribution to the peer review of this work.

**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.

**Acknowledgements**
We are highly grateful to Dr. Nikolay Korolev for valuable discussions. We are highly grateful to Prof. Kristian Helin and Dr. Paul Cloos for critical reading of the manuscript. We are thankful to Dr. Heike Ilona Könnar for helpful discussions in the early stages of the project. The plasmid constructs for 15-197-601, 16-187-601, and 12-202-601 DNA were kindly provided by Prof. Daniela Rhodes. We are thankful to Jens Vilsstrup Jøhansen at BRIC bioinformatic core facility for pre-processing of the ChIP-seq and ATAC-seq datasets. Work in the CSS laboratory was supported by the Novo Nordisk Foundation, the Danish Cancer Society, and the Danish National Research Foundation for grant DNRF115 funding the Center for Chromosome Stability. The Singapore Ministry of Education Academic Research Fund (AcRF) Tier 2 (MOE2018-T2-1-112) and Tier 1 (2018-T1-001-114) grants supported work in the LN laboratory. All ssNMR experiments were performed at the Nanyang Technological University (NTU) Center of High-Field NMR Spectroscopy and Imaging. We also acknowledge the NTU Institute of Structural Biology (NISB) for supporting this research. ChIP-sequencing was performed at the Bioinformatics and Expression Analysis (BEA) core facility at KI, which is supported by the board of research at the Karolinska Institutet and the research committee at the Karolinska Hospital, ChIP samples were prepared by Christos Coulcaravas. Work in the KE laboratory was supported by the Swedish Research Council and Cancerfonden.