Oct4 interacts with subnucleosomal particles generated by Brg1 at enhancers

Marina C. Nocente[1,4], Anida Mesihovic Karamitsos[1,4], Emilie Drouineau[1,4], Waad Albawardi[2], Cécile Dulary[3], Florence Ribierre[3], Hélène Picaud[1], Olivier Alibert[3], Jean-Christophe Aude[1], Nick Gilbert[2], Françoise Ochsenbein[1], Sophie Chantalat[3], Matthieu Gérard[1]

1. Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France
2. MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, Crewe Road South, Edinburgh, EH4 2XU, UK
3. Institut de biologie François Jacob, Centre National de Recherche en Génomique Humaine, 91057 Evry, France
4. These authors contributed equally

Contact : matthieu.gerard@i2bc.paris-saclay.fr

ABSTRACT

Mammalian SWI/SNF complexes (mSWI/SNF) are ATP-dependent chromatin remodelling factors that possess tumour suppressor properties[1]. Brg1 (Smarca4), the catalytic subunit of mSWI/SNF, is essential for chromatin opening at enhancers[2,3], but the detailed mechanism of this activity remains unclear. Using mouse embryonic stem (ES) cells, we show that Brg1 generates subnucleosomal particles which protect genomic DNA fragments of 50-80 bp in length from micrococcal nuclease (MNase) digestion. These particles, which contain all four histones H3, H4, H2A, and H2B, sediment in sucrose gradients with an apparent molecular weight markedly inferior to that of canonical nucleosomes. Genome-wide analysis of the positioning of these particles provides strong evidence that they arise from the splitting of Brg1-targeted nucleosomes into hemisomes. *In vivo*, these hemisome-like particles are bound by Oct4 (Pou5f1), a transcription factor essential for the maintenance of the ES cell phenotype. Oct4 has a mode of interaction with these particles that is distinct from its binding to accessible, histone-free DNA. Whereas Oct4 binds the latter substrate at the level of its consensus DNA motif, its interaction with the subnucleosome persists even when Oct4 motifs are removed from the particle, suggesting that Oct4 binding is stabilized by contacts with histones. Our data establish that one critical activity of Brg1 at enhancers involves splitting selected nucleosomes into subnucleosomes, which are major genomic binding targets for transcription factors.

MAIN TEXT

In ES cells, enhancers are bound by a series of transcription factors (TFs) that are required to maintain ES cell pluripotency and self-renewal, including Oct4, Sox2, Nanog, Esrrb and Klf4[4]. The DNA binding sites for these TFs are located in relative proximity to each other, yet without a constrained arrangement pattern[4–8]. Enhancers are also characterized by their hypersensitivity to DNase I, the presence of the mediator complex, and, for active enhancers, RNA polymerase II (Pol II) and TATA box Binding Protein (TBP), as well as H3K27ac[9–12]. However, the chromatin organization of enhancers remains less defined than that of promoters and CCCTC-binding factor (CTCF)-bound loci[13–20]. This work investigates nucleosomal organization at enhancers by micrococcal nuclease digestion of chromatin, followed by immunoprecipitation using antibodies against histone H3 and deep sequencing (MNase ChIP-seq) (**Fig. 1a**). We show that k-means clustering based on the distribution patterns of canonical nucleosomes and subnucleosomal-size particles leads to identifying clusters of enhancers with different but highly related chromatin organizations (Extended Data Fig. 1a).

Using two-dimensional maps (V-plot)[21], we detected in the central region of each cluster of enhancers the presence of short (50-120 bp) DNA fragments associated with histone H3, corresponding to subnucleosomal particles precisely positioned across the enhancers (Extended Data Fig. 1b). These subnucleosomal particles alternate in position with one or more full-size nucleosomes (i.e., histone-associated 150-180 bp long DNA fragments), depending on the distance that separates the two well-positioned nucleosomes, defined hereafter as nucleosomes -1 and +1, that flank the centre of the enhancer (**Fig. 1b** and Extended Data Figs. 1b, 2a). Enhancer clusters 1 to 3 that are representative of all subgroups differ from each other according to the distance between nucleosomes -1 and +1, with cluster 1 having the narrowest (340 bp) and cluster 3 the widest (> 480 bp) distance (Extended Data Fig. 1b).

In enhancer cluster 1 and related clusters 5, 8 and 12, a single centrally located nucleosome is present between -1 and +1 nucleosomes (**Fig. 1b** and Extended Data Fig. 1b). This central nucleosome is highly sensitive to MNase dosage, as it completely disappeared when ES cells were incubated with a four-fold excess of MNase (**Fig. 1c**), and we will hereafter refer to this particle as a fragile nucleosome[15]. In contrast, the flanking nucleosomes -1 and +1 were only slightly affected by the high-dose MNase treatment, revealing their canonical nature.

Enhancer cluster 2, and related clusters 4, 9 and 10, are characterized by the presence of fragile nucleosomes at two main positions instead of one in the central interval between nucleosomes -1 and +1 (Extended Data Fig. 2a, b). We observed that in cluster 1, the single central fragile nucleosome is symmetrically flanked on each side of its dyad axis by a subnucleosomal particle protecting from MNase digestion a minimum DNA length of 54 bp, as assessed by the length of the DNA fragments present at the vertex of the 'V' shape signal (**Fig. 1b** and Extended Data **Fig. 3a**). In cluster 2, each of the two fragile nucleosomes of the central interval is flanked by subnucleosomal particles in a pattern similar to cluster 1 (Extended Data Fig. 2a and Extended Data Fig. 3a). Enhancer cluster 3 shows a variation of these patterns, with a wider central interval accommodating a larger number of subnucleosomal particles (Extended Data Fig. 1b). Thus, our analysis revealed that enhancers have a modular chromatin organization in which two positioned canonical nucleosomes delimit a central region containing a variable number of basic modules. Each basic module is composed of one fragile nucleosome flanked on both sides of its dyad axis by two subnucleosomal particles occupying precisely the same genomic interval, suggesting that they are alternating chromatin configurations.

To investigate the relation between this discrete chromatin architecture and TF motif distribution, we mapped the DNA-interacting motifs of the pluripotency-associated TFs Oct4, Sox2, Nanog, and Klf4[4]. The genomic interval encompassing the central fragile nucleosome(s) and associated-subnucleosomal particles corresponds precisely to the area in which TF motifs are enriched (Extended data Fig. 3b, c). Thus, the larger size of this genomic interval at multi-modular enhancers correlates with a greater number of Oct4 and other TF motifs compared to single-module enhancers (Extended data Fig. 3). We then examined the distribution of TFs and epigenetic features that contribute to enhancer functions (**Fig. 1g-l** and Extended Data Fig. 2f-l). ChIP-seq analysis revealed that all tested TFs bind within the central interval of the enhancer in which subnucleosomal particles are interspaced with fragile nucleosome(s). Even TFs known to have pioneer functions, such as Oct4 and Sox2, were preferentially enriched on DNA fragments spanning 50 to 120 bp in length rather than the 150-180 bp typical of canonical nucleosomes (**Fig. 1g, h** and Extended Data Fig. 2f, g). The general transcription factor TBP and Pol II were also detected within the same central interval of the enhancer and were associated with 40-120 bp DNA fragments (**Figure 1k, l** and Extended Data Fig. 2i, k). In contrast to nucleosomes -1 and +1, the subnucleosomal particles were unexpectedly devoid of H3K27ac signal (**Fig. 1j** and Extended Data Fig. 2j). The

enhancers selected for this study lack CTCF ChIP-seq signal (Methods), excluding the possibility that this TF might organize nucleosome positioning.

We then investigated which factors control this enhancer-specific nucleosomal organization. ATP-dependent chromatin remodellers[22–25] are prime candidates; in particular, Brg1 (Smarca4), the ATPase of mammalian SWI/SNF complexes[26], is required for the generation of accessible chromatin at enhancers[2,3,27]. Depletion of Brg1 using an shRNA resulted in striking alterations of both nucleosomal and subnucleosomal organization at enhancers. Subnucleosomal particles were almost completely absent at enhancers in cells depleted of Brg1, compared to control cells (**Fig. 1e** and Extended Data Fig. 2d). In addition, the positioning of canonical nucleosomes was severely altered, resulting in their apparently random redistribution across the enhancer. We detected identical alterations with antibodies against histone H3 and H2B, as well as when we used a second shRNA targeting a different region of Brg1 mRNA (Extended Data Fig. 4a). These perturbations of chromatin organization at enhancers were specific to Brg1 loss-of-function, as shRNA-mediated depletion of chromatin remodeller Chd4 did not prevent the generation of subnucleosomal particles, nor the positioning of canonical nucleosomes (**Fig. 1f** and Extended Data Fig. 2e). Depletion of Brg1 using the auxin-induced degron (AID) system[28] resulted in similar defects in chromatin organization (Extended Data Fig. 4b, c).

The alterations in chromatin organization caused by Brg1 loss-of-function are compatible with a scenario in which Brg1 binds the central fragile nucleosome(s) to convert it into smaller particles, such as hexasomes [$(H3\text{-}H4)_2$; H2A-H2B] or hemisomes (half-nucleosomes) containing one copy of each of the histones H3, H4, H2A, H3B[29–32]. In support of this hypothesis, we detected Brg1 enriched both at the centrally located fragile nucleosome(s) and at subnucleosomal particles (**Fig. 1d** and Extended Data Fig. 2c). In an alternative hypothesis, Brg1 could extrude a DNA loop on the central nucleosome and propagate it to the dyad axis region[33], which would become highly sensitive to MNase digestion and yield a histone octasome associated with two short DNA fragments symmetrically positioned on each side of the dyad, instead of the regular 150-180 bp-long fragment. To test these hypotheses, we set up an experimental strategy based on separating nucleosomal and subnucleosomal particles by centrifugation through a sucrose gradient, in which the particles sediment as a function of their molecular mass and shape[34] (**Fig. 2a**). MNase-digested ES cell chromatin was loaded onto a 10-30% sucrose gradient. After centrifugation, individual fractions were collected and analysed by agarose gel electrophoresis, revealing genomic DNA fragments ranging from 30 to 120 bp, which

migrated slower than canonical nucleosomes through the gradient (**Fig. 2b**). To identify which fractions contained subnucleosomal particles, we performed ChIP-seq experiments with an antibody against histone H3 (**Fig. 2c**). At the top of the gradient, the DNA fragments in fractions 3-4 were not efficiently immunoprecipitated, suggesting that they mostly correspond to histone-free DNA fragments. In contrast, the 50-80 bp-long particles collected in gradient fractions 5-6, which have a low sedimentation rate, were associated with histone H3 at enhancers and are thus genuine subnucleosomal particles (**Fig. 2c**). Their slow sedimentation rate is incompatible with the hypothesis that these particles correspond to histone octamers associated with two 50-80 bp DNA fragments. Such particles would have approximately the same molecular weight as canonical nucleosomes and thus would be expected to sediment in fractions 11 to 13.

To define the histone content of the low sedimentation rate 50-80 bp subnucleosomal particles (hereafter called 50-80 bp subnucleosomes), we performed ChIP-seq experiments with antibodies against the remaining three histones H4, H2A and H2B. These antibodies efficiently immunoprecipitated the 50-80 bp subnucleosomes but not the DNA fragments from the top of the gradient (**Fig. 3**). This result excludes the possibility that the 50-80 bp subnucleosomes might correspond to tetrasomes composed of two dimers of histones H3 and H4, which are nucleosome assembly intermediates detected during DNA replication[35]. The presence of the four histones suggests that these particles correspond to either hexasomes or hemisomes[29,31,32,36]. However, the DNA length (50-80 bp) protected from MNase digestion is not compatible with the ~90 bp of DNA associated with hexasomes[37]. The protection from MNase cleavage conferred by hexasomes yielded DNA fragments ranging from 90 to 130 bp[29]. In our experimental setting, subnucleosomal particles associated with 90-130 bp DNA fragments have a medium sedimentation rate (fractions 7-8 and 9-10), suggesting a histone composition different from the 50-80 bp subnucleosomes that sedimented in fractions 5-6 (**Fig. 2**). Furthermore, the positions of the DNA fragment centres of these 50-80 bp particles are symmetrically distributed on each side of the dyad axis of the central fragile nucleosome, and the distance spanning the sum of the lengths of the two consecutive particles (140-180 bp) is equivalent to the size of DNA wrapped around an octasome (Extended Data Fig. 3a). Examination of individual single-module enhancers from clusters 1 and 8 revealed that 50-80 bp subnucleosomes are systematically associated in pairs, and separated by a DNA linker ranging in size from 0 to 102 bp (median 29 bp) (**Fig. 4d**, Extended Data Fig. 5 and Methods). This unique positioning of the 50-80 bp subnucleosomes, their histone content, and their low sedimentation rate suggest that they correspond to split nucleosomes that are composed of

two hemisomes connected by a linker DNA fragment of variable length. This linker DNA is likely hypersensitive to MNase digestion, which separates the two hemisomes from each other and results in their co-segregation within the low sedimentation rate fractions of the gradient.

We took advantage of our high-coverage histone MNase ChIP-seq datasets to explore subnucleosomal organization at CTCF-binding sites and promoters. This analysis revealed the presence of 50-80 bp subnucleosomes having distribution patterns specific to each category of cis-regulatory elements (Extended Data Fig. 6). At CTCF-binding sites, a single low sedimentation rate particle was detected in the central region coinciding with the location of the CTCF motif (Extended Data Fig. 6a, b). Promoters also contain a single low sedimentation rate particle rather than the paired particle pattern of enhancers, coinciding in genomic position with nucleosome -1 located upstream of the TSS (Extended Data Fig. 6c, d). This observation suggests that at promoters, 50-80 bp subnucleosomes are the products of remodelling events targeting nucleosome -1, which is bound by a large variety of chromatin remodellers in ES cells[23]. This nucleosome was hypersensitive to an excess of MNase (Extended Data Fig. 6c), a hallmark of fragile nucleosomes. We detected a 30 bp protection against MNase digestion (footprint) within the 50-80 bp subnucleosomes of promoters, revealing the binding of proteins potentially involved in preinitiation complex formation (Extended Data Fig. 6e). Since these particles are present in low sedimentation rate fractions, the footprint might be conferred by low molecular weight TFs like TBP and TFIIB but not by large complexes such as TFIID or Pol II.

The chromatin perturbation pattern caused by Brg1 depletion (**Fig. 1**) suggests that this remodeller might generate the 50-80 bp subnucleosomes at enhancers. We tested this possibility by repeating the sucrose gradient sedimentation experiments, using chromatin prepared this time from either Brg1-depleted or control cells. ChIP-seq experiments with antibodies against histones H3, H4, H2A and H2B revealed that Brg1 depletion results in a strong reduction of the amount of 50-80 bp subnucleosomes, thus indicating that their production at enhancers is indeed dependent on Brg1 (**Fig. 3**). Brg1 depletion also reduced the amount of medium sedimentation rate 90-130 bp subnucleosomal particles, showing that Brg1 activity is required for the generation of all classes of subnucleosomal particles detected at enhancers in this study (Extended Data Fig. 7c-f). In contrast, Brg1 depletion did not alter the distribution or amount of subnucleosomes at CTCF binding sites (Extended Data Fig. 6b). At promoters, Brg1 loss of function did not affect nucleosome positioning, but the amount of 50-80 bp subnucleosomes was slightly decreased (Extended Data Fig. 6c, d). Altogether, these

results show that Brg1 activity is mainly required at enhancers to produce subnucleosomal particles and properly position nucleosomes.

We also tested whether Brg1 is involved in producing histone-free DNA fragments at enhancers. Deep sequencing of the DNA prepared from the top gradient fractions revealed that Brg1-depletion resulted in a pronounced reduction of these DNA fragments at enhancers but not at CTCF-binding sites (Extended Data Fig. 7a, b). DNA fragments containing motifs for Oct4, Sox2, Esrrb, Klf4 and Nanog were markedly depleted in the top gradient fractions prepared from Brg1-depleted cells (Extended Data Fig. 7g), demonstrating that Brg1 is required for the efficient production of histone-free, accessible DNA for these TFs at enhancers.

We observed that in the Oct4, Sox2 and Nanog ChIP-seq experiments (**Fig. 1g-i**), the sizes and position on the genome of the immunoprecipitated DNA fragments were similar to that of 50-80 bp subnucleosomes (**Fig. 3**). This specific pattern, reproduced with three distinct antibodies against Oct4 (**Fig. 1, 4** and Methods), suggests that 50-80 bp subnucleosomes might represent a previously undetected binding substrate for TFs *in vivo*. To test if Oct4 is genuinely bound to histone-associated subnucleosomal particles, we carried out sequential ChIP experiments using antibodies against Oct4 in the first step. Elution of Oct4-bound chromatin was performed using a competitor peptide, and the eluted chromatin was used in a second step of ChIP with antibodies against histone H2B. Deep sequencing of the DNA associated with both Oct4 and histone H2B revealed DNA fragments of subnucleosomal size, thus demonstrating that Oct4 is bound to subnucleosomal particles *in vivo* (**Fig. 4e**).

We next compared how Oct4 binds to 50-80 bp subnucleosomes and histone-free DNA present in the top gradient fractions. Oct4 antibodies successfully immunoprecipitated both histone-free DNA fragments and 50-80 bp subnucleosomes (**Fig. 4a, b**). However, Oct4 enrichment profiles were different at these two distinct genomic templates. At histone-free DNA fragments, Oct4 peaks coincided with Oct4 consensus motifs, whereas in 50-80 bp subnucleosomes, Oct4 enrichment coincided with histones (**Fig. 4d** and Extended Data Fig. 5, 8). The specificity of this unexpected Oct4 ChIP-seq enrichment profile on 50-80 bp subnucleosomes was verified with a second antibody against Oct4, and by performing the elution step of the ChIP experiment using a competitor peptide (**Fig. 4b-d**, and Extended Data Fig. 5, 8). We observed this coincidence between Oct4 and histone enrichment even when Oct4 binding motifs were absent from the DNA wrapped around the particle (Extended Data Fig. 9). This apparent paradox suggests that Oct4 interactions with 50-80 bp subnucleosomes involve contacts with the histone component in addition to those with its DNA binding motif.

Oct4 binding motifs are strongly enriched in the dyad region of the central fragile nucleosome of the enhancer (Extended Data Fig. 3b). Brg1-mediated conversion of this central nucleosome into two hemisome-like particles is associated with the appearance of a DNA linker between the two particles. This linker contains the Oct4 motifs initially present in the dyad region of the central nucleosome (Extended Data Fig. 5i). MNase treatment prior to ChIP-seq cuts away the linker; Oct4 would be released if it interacted with the subnucleosomal particle exclusively via its DNA motif. Since Oct4 remains bound to hemisome-like particles that do not contain an Oct4 motif (Extended Data Fig. 9), we concluded that these particles must provide an interface stabilising Oct4 interaction. An intriguing possibility is that the internal histone surface of the split nucleosome, which is inaccessible in canonical nucleosomes, provides an interaction domain for transcription factors and potentially for other chromatin-binding proteins in a manner reminiscent of the acidic patch of the nucleosome, which is present on the other side (**Fig. 4f**).

We next used transmission electron microscopy (EM) to compare the size of 50-80 bp subnucleosomes with both mononucleosomes and nucleosomes assembled *in vitro* onto an array of 601 DNA fragments[38] (ED Fig. 10a, b). Canonical nucleosomes have a front view diameter of 10 to 13 nm and a width of about 6 nm, the latter of which is defined by the sum of the two DNA gyres wrapped around the histone octamer[39]. If the particles generated at enhancers are hemisomes, they should retain a 10-13 nm diameter but have a width reduced to 3 nm, which corresponds to the size of a single DNA gyre. EM analysis should then reveal particles having an apparent size either smaller or equal to that of canonical nucleosomes, depending on their orientation on the grid. Random selection and size measurement of both particles from gradient fractions 5-6 and particles further immunopurified with Oct4 antibodies revealed a large proportion of small-sized particles (< 10 nm), which are absent in control nucleosome-containing fractions (ED Fig. 10a, b). The EM analysis thus further supports the hypothesis that enhancer 50-80 bp subnucleosomes correspond to hemisomes.

Finally, we investigated whether the 50-80 bp subnucleosomes are also present in human cells. We used the histone MNase ChIP-seq approach combined with sucrose gradient sedimentation experiments to probe chromatin organization in a melanoma cell line. We isolated clusters of putative enhancers enriched for 50-80 bp subnucleosomes, which have a chromatin organization very similar to that of mouse ES cell enhancer clusters (Extended Data Fig. 10c-e). This analysis shows that the modular organization of nucleosomes and subnucleosomes described at mouse ES cell enhancers is conserved in other mammalian cell types.

This study establishes mSWI/SNF as a chromatin remodeller dedicated to generating 50-80 bp subnucleosomes at enhancers. Our results provide strong evidence that these subnucleosomes are hemisomes resulting from the splitting of fragile nucleosomes mediated by Brg1 (**Fig. 4f**). The specific mode of interaction of Oct4 with these particles indicates that they are not mere intermediates in nucleosome remodelling events leading to histone free DNA. These subnucleosomes appear to correspond to a chromatin template stabilising TF binding at enhancers. Since we show that related particles are also generated at promoters, such TF-subnucleosome interactions might have an important function in regulating transcription *in vivo*.

FIGURES

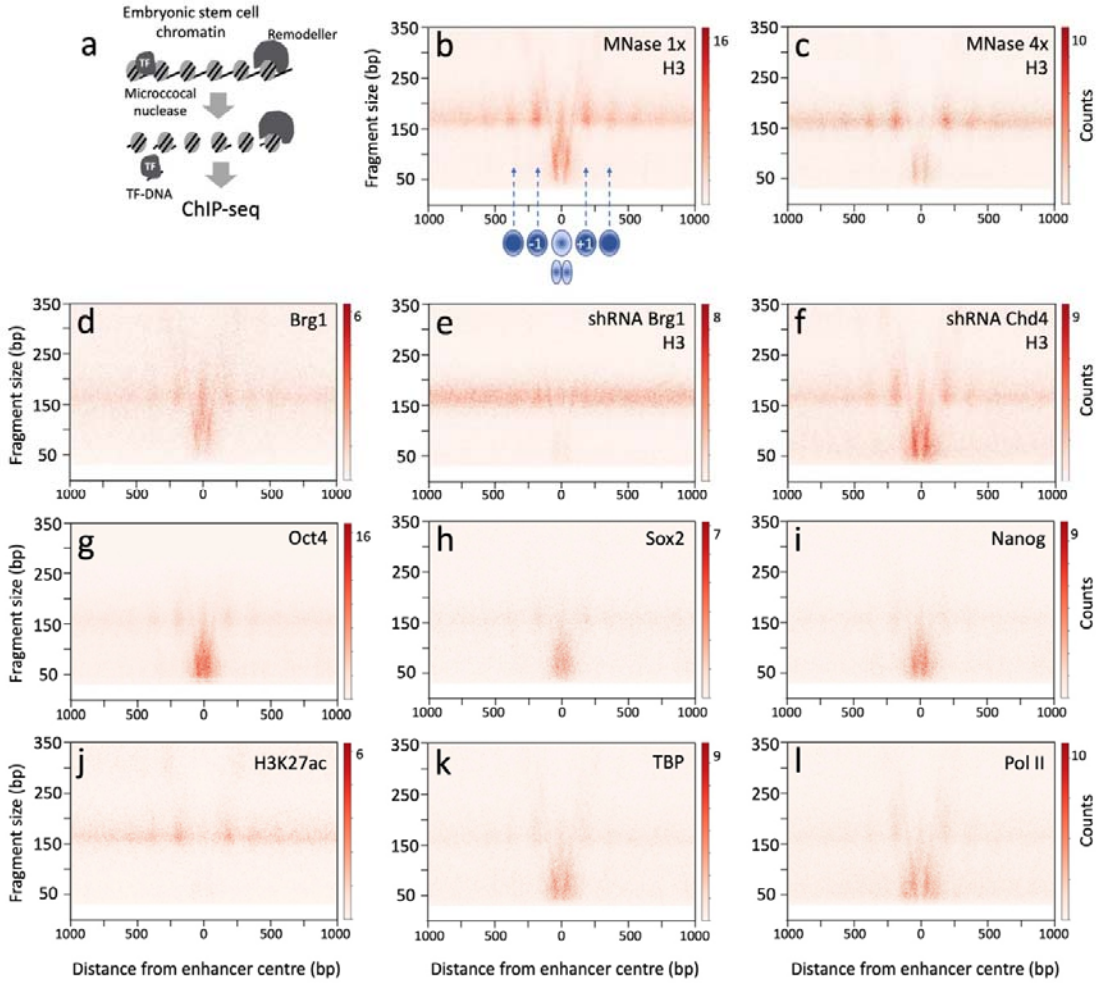


Figure 1. Brg1 controls nucleosomal and subnucleosomal organization at enhancers.

a, MNase-digested chromatin was prepared from ES cells for ChIP-seq experiments. **b**, **c**, V-plots of histone H3 ChIP-seq fragments spanning ± 1000 bp from cluster 1 enhancer centre, using either the standard MNase dose (**b**), or a four-fold excess (**c**). Red dots indicate on the *x-axis* the genomic position of the midpoint of each immunoprecipitated DNA fragment and its length in bp on the *y-axis*. The colour scale corresponds to the number of DNA fragments. The schematic illustration in (**b**) indicates the positions of nucleosomes -1 and +1, which flank the central fragile nucleosome and associated subnucleosomal particles. **d,** V-plot of Brg1 ChIP-seq. **e**, **f**, V-plots of histone H3 ChIP-seq in ES cells depleted of either Brg1 (**e**), or Chd4 (**f**), using shRNAs. **g-l**, V-plots of Oct4, Sox2, Nanog, H3K27ac, TBP and Pol II ChIP-seq. The standard MNase dose was used in all panels except in (**c**).

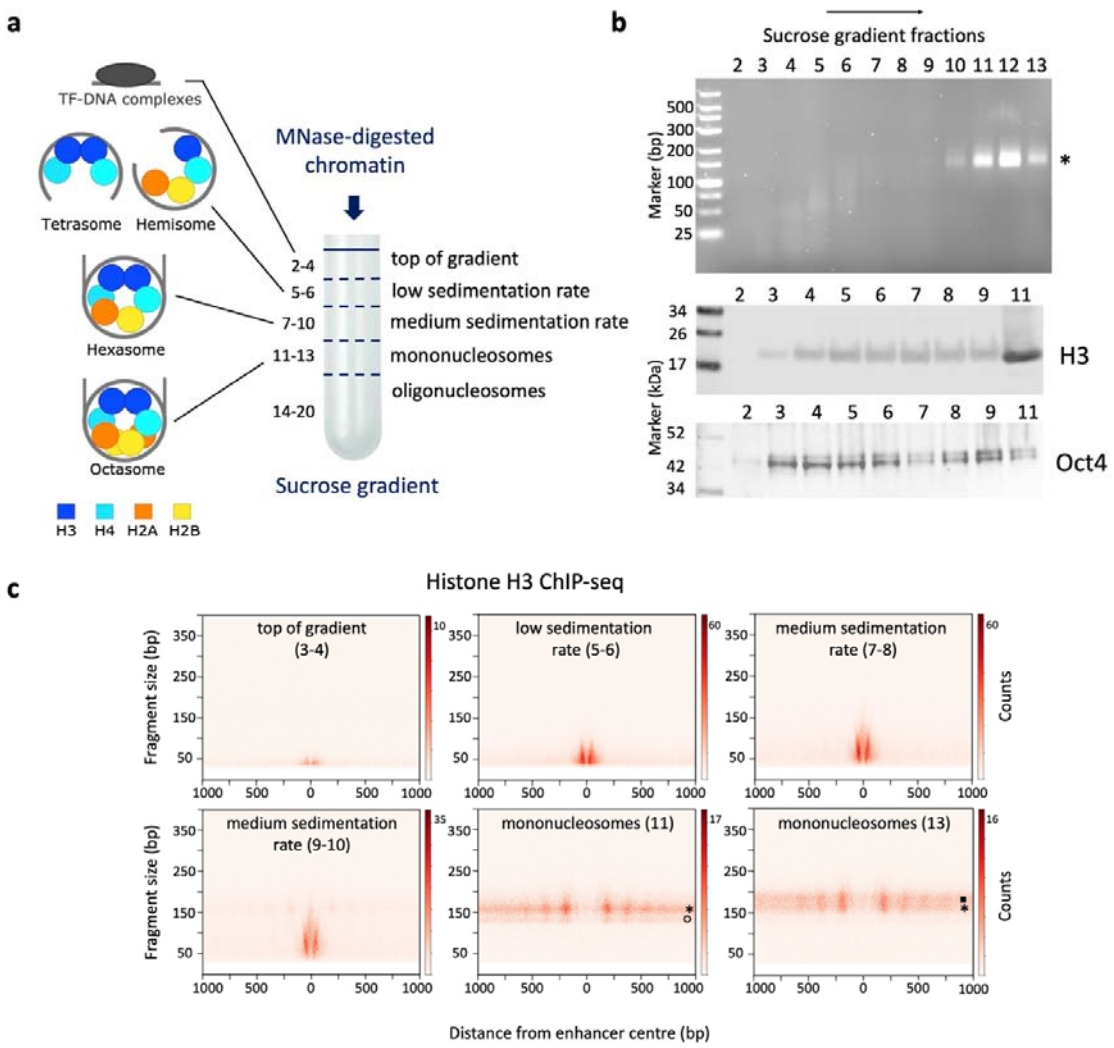


Figure 2. Identification of low sedimentation rate subnucleosomal particles encompassing 50-80 bp of DNA at enhancers. a, MNase-digested chromatin was centrifugated through a 10-30% sucrose gradient to separate canonical nucleosomes from putative subnucleosomal particles and TF-DNA complexes. **b**, The DNA fragments of each fraction were analysed by electrophoresis in a 4% agarose gel (top), and the distribution of histone H3 and Oct4 was revealed by Western blot (mid and bottom). The asterisk indicates the DNA fragments corresponding to mononucleosomes. **c**, V-plots of histone H3 ChIP-seq experiments performed with sucrose gradient fractions as input, spanning ± 1000 bp from cluster 1 enhancer centre. The fractions or pools of fractions used for ChIP are indicated on each panel. The open circle and black square point to potential partially unwrapped nucleosomes and chromatosomes, respectively.

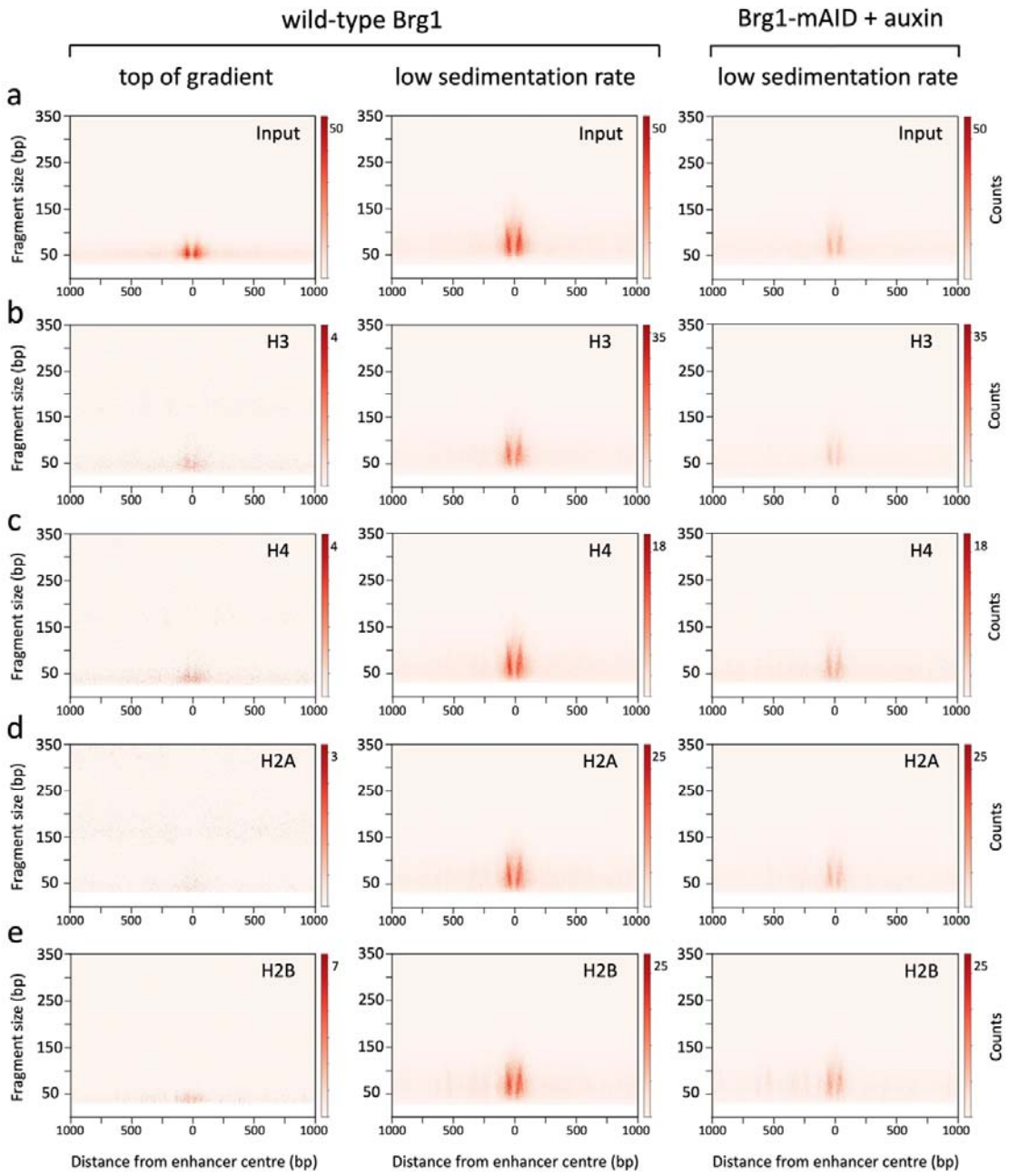


Figure 3. Histone composition of 50-80 bp subnucleosomes. Chromatin, prepared from ES cells expressing wild-type Brg1, or from cells depleted of Brg1 using the AID system, was centrifugated through a sucrose gradient as in Fig. 2. **a**, V-plots of the DNA fragments which sedimented at the top of the gradient (pool of fractions 2-3-4), or of the DNA fragments present in low sedimentation rate fractions (pool of fractions 5-6), spanning ± 1000 bp from the cluster 1 enhancer centre. Signals shown in the two right panels were adjusted according to the total number of fragments sequenced in each dataset. **b-e**, V-plots of histone H3, H4, H2A and H2B ChIP-seq experiments performed with the pools of sucrose gradient fractions indicated in (**a**) as input. For each histone, the signals of the ChIP-seq experiments performed with subnucleosomal particles from ES cells expressing wild-type Brg1

and from Brg1-depleted cells were adjusted according to the total number of DNA fragments in each dataset.

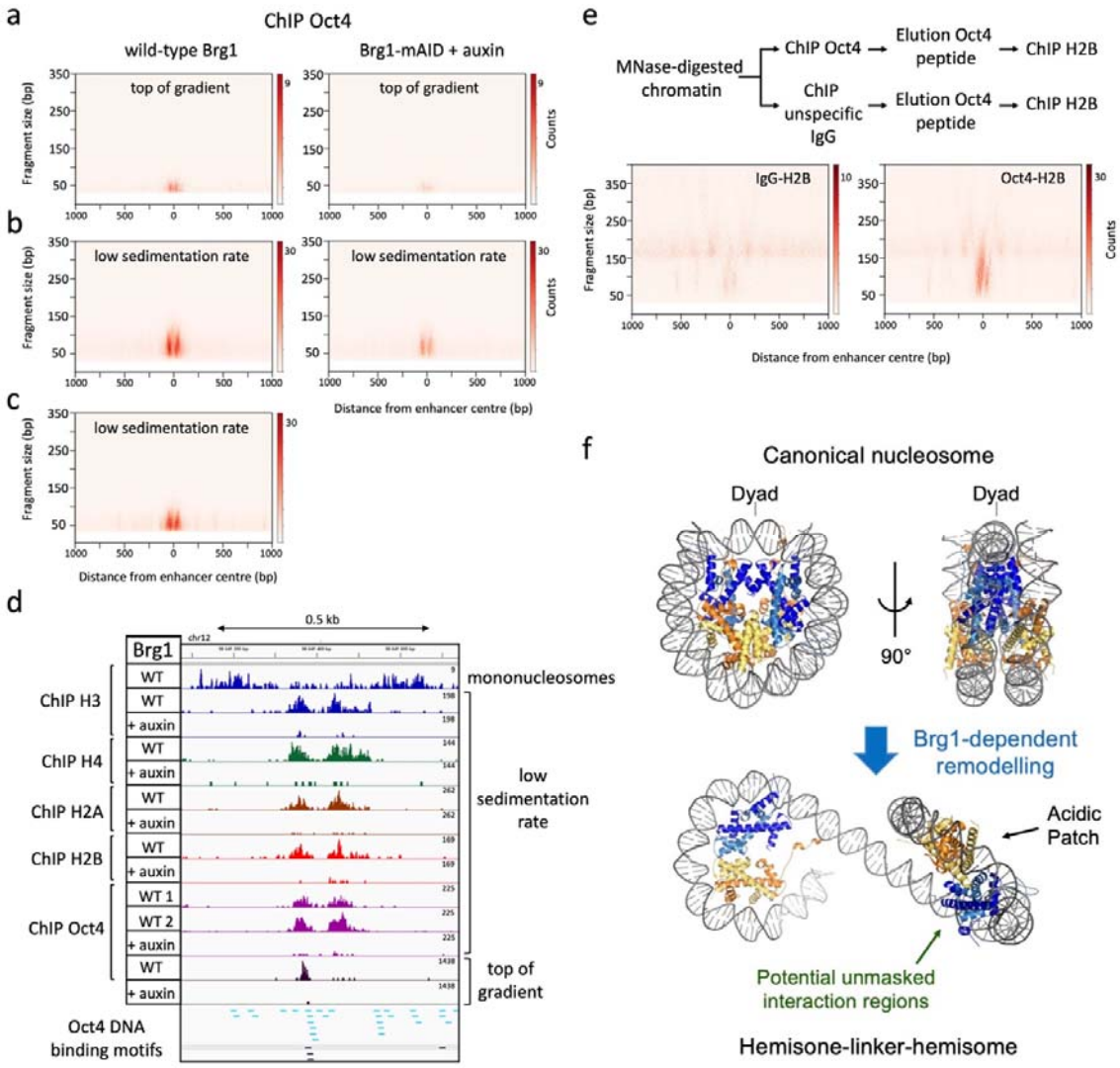
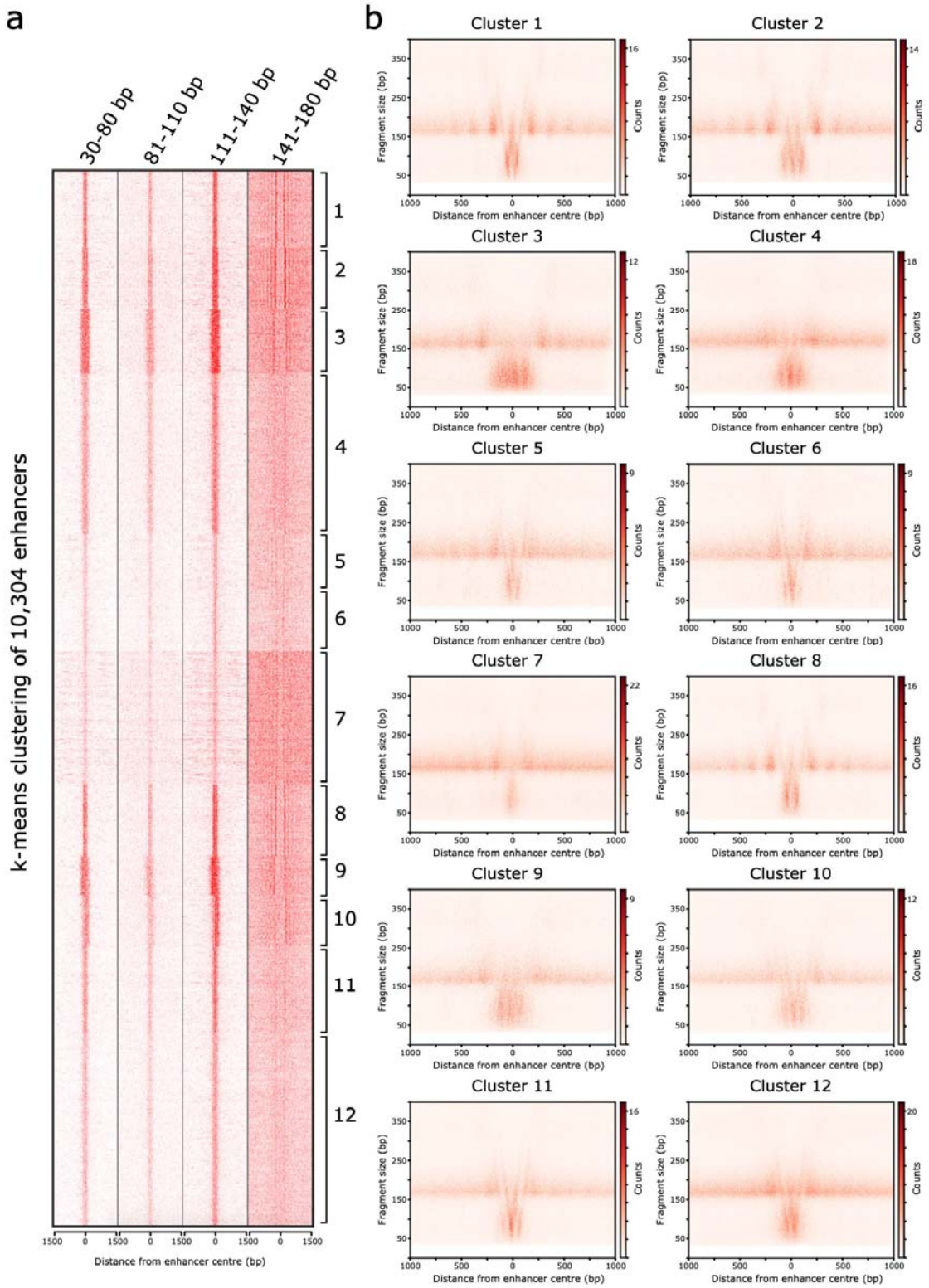


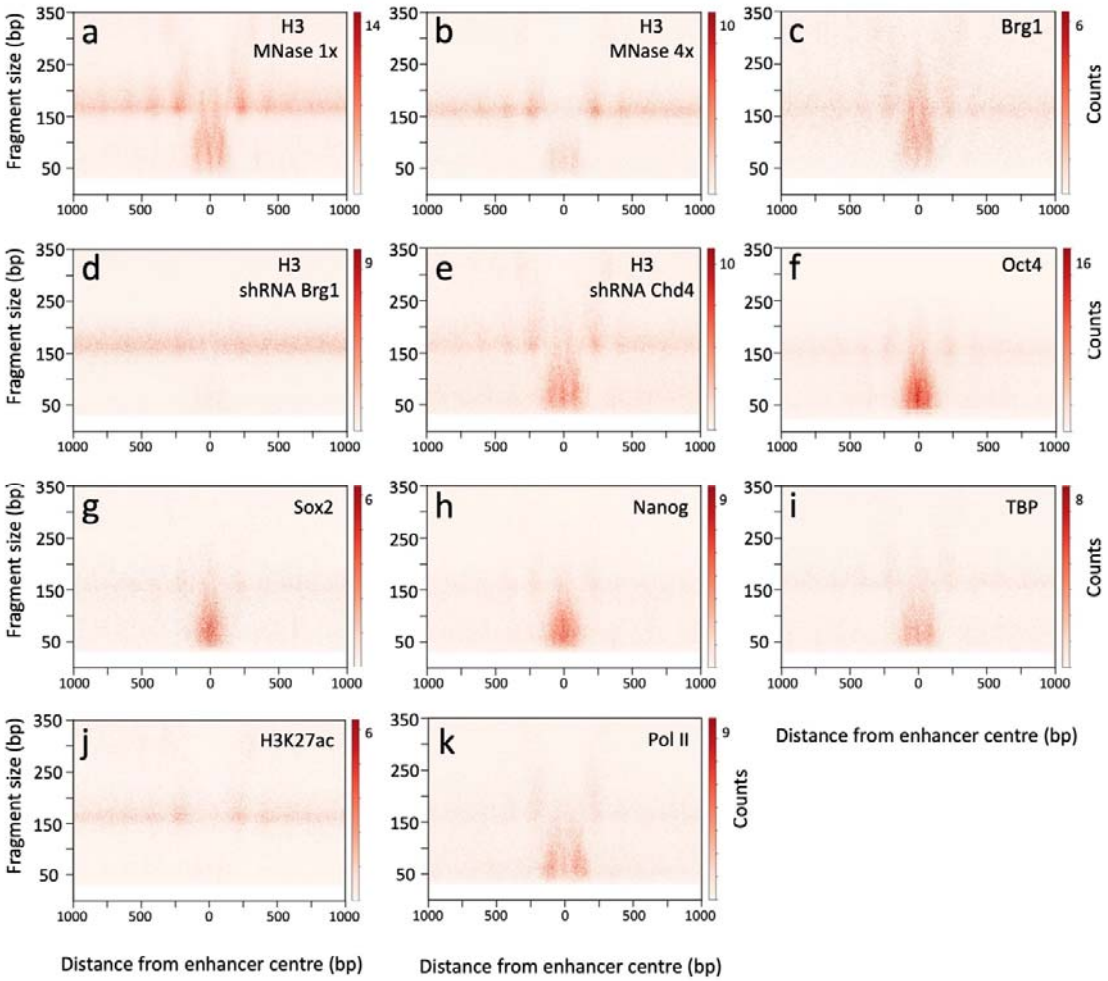
Figure 4. Oct4 binds to 50-80 bp subnucleosomes at enhancers. a-c, V-plots of ChIP-seq fragments spanning ± 1000 bp from the cluster 1 enhancer centre. Chromatin, prepared from mouse ES cells expressing wild-type Brg1 (left panels) or from cells depleted of Brg1 using the AID system (right panels), was centrifugated through a sucrose gradient as in Fig. 2. Oct4 ChIP experiments were performed using DNA and chromatin fragments which sedimented at the top of the gradient (**a**, pool of fractions 2-3-4), or present in low sedimentation rate fractions (**b**, **c**, pool of fractions 5-6). Two distinct Oct4 antibodies were used in (**b**) and (**c**). **d**, Density graphs showing the distribution of ChIP DNA fragment centres at a representative example of a cluster 1 enhancer. The top lane shows the positions of canonical nucleosomes detected by histone H3 ChIP-seq of sucrose gradient fractions 11 and 13. The next eight lanes show the distribution of 50-80 bp subnucleosomes, revealed by histone H3, H4, H2A and H2B ChIP-seq. The following five lanes show the Oct4 ChIP-seq signal on 50-80 bp subnucleosomes or histone-free DNA from the top gradient fractions. Experiments were performed with chromatin of cells expressing wild-type (WT) Brg1 or depleted of Brg1 using the AID system (+

auxin). Two distinct antibodies against Oct4 were used in lanes WT 1 and 2. High ($p = 0.001$) and low ($p = 0.01$) confidence Oct4 consensus binding motifs are indicated in black and blue, respectively. **e**, Sequential ChIP-seq experiments were performed on chromatin prepared from ES cells expressing wild-type Brg1. Chromatin was first immunoprecipitated with an Oct4 antibody or with an unspecific IgG. Elution was performed by peptide competition, and each eluted fraction was subjected to a second round of immunoprecipitation with an antibody against H2B. The right and left panels show the sequential Oct4-H2B ChIP-seq and the unspecific IgG control experiment. **f**, Model proposing that Brg1 splits nucleosomes at enhancers to generate hemisome-linker-hemisome entities, which might display new interaction regions for TFs. A linker length of 0 to 102 bp (median = 29 bp, shown in this figure) separates the two hemisomes. Histones H3 and H4 are shown in two shades of blue, and histones H2A and H2B in yellow and orange, respectively.

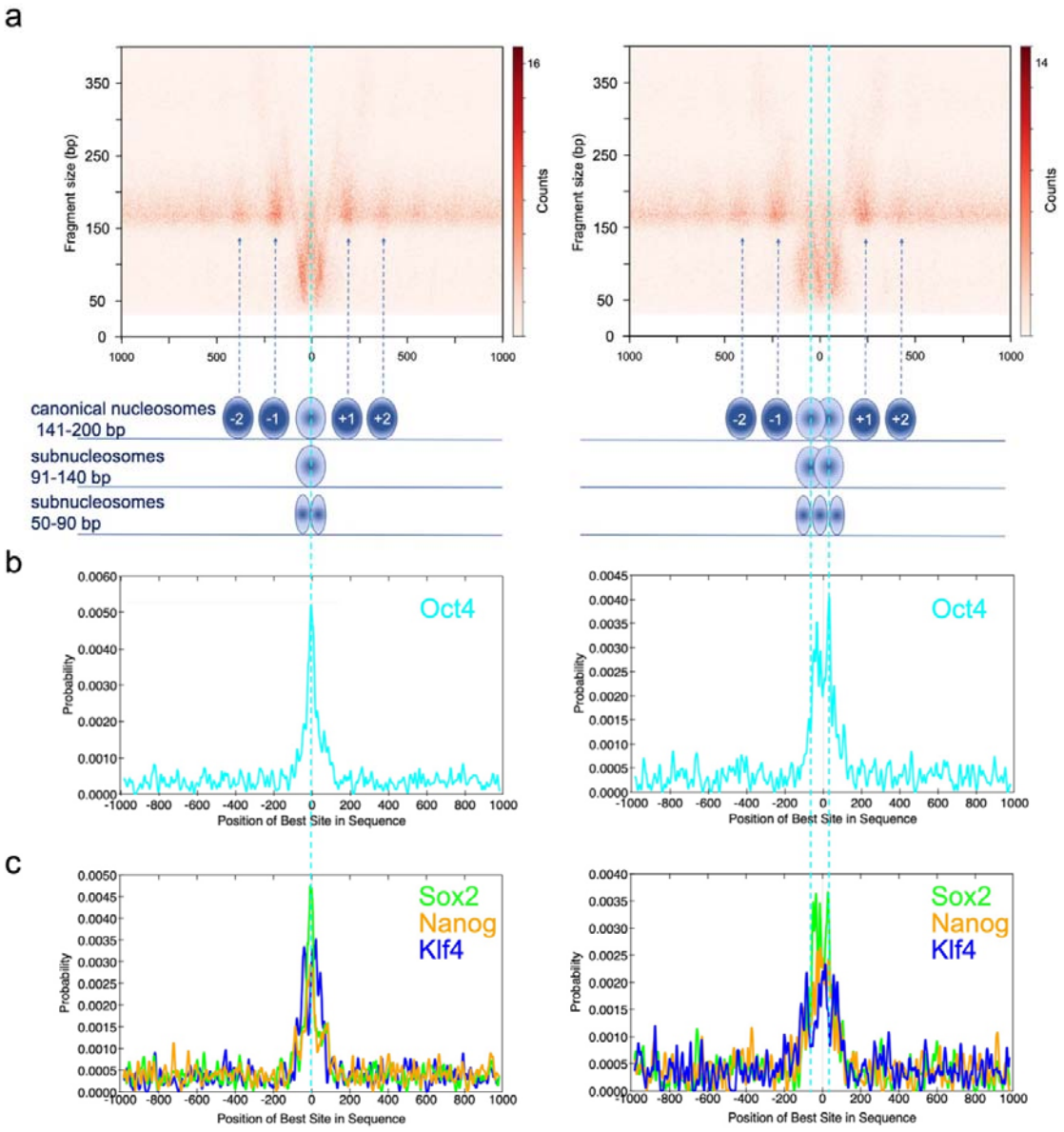


Extended Data Figure 1. Integrated nucleosomal and subnucleosomal organization at enhancers. a, Heat map of histone H3 enrichment at 10,304 enhancers determined by MNase-ChIP-seq (left panel). The four columns display histone H3-immunoprecipitated DNA fragments of different

size categories. Enhancers were separated into 12 groups by k-means clustering. The enhancer centre corresponds to the middle of its DNase I hypersensitivity peak. **b**, For each group, the distribution of histone H3 ChIP-seq fragments was analysed by two-dimensional maps (V-plot), showing the distribution and density of DNA fragments immunoprecipitated with antibodies as a function of the position of their mid-point on the genome (x-axis), and their length (y-axis). The colour scale corresponds to the number of DNA fragments.

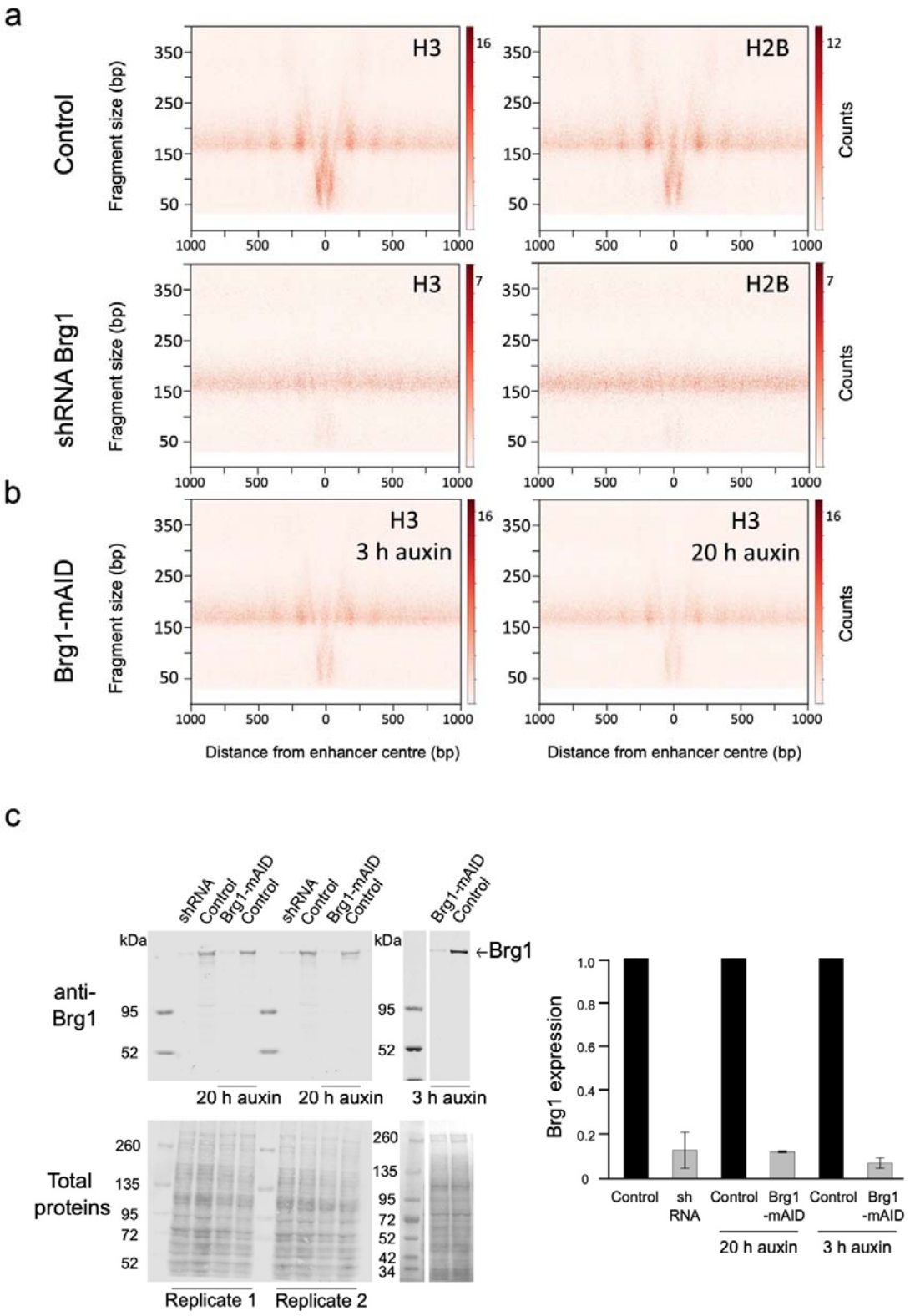


Extended Data Figure 2. Nucleosomal and subnucleosomal organization of cluster 2 enhancers. a-b, V-plots of histone H3 ChIP-seq fragments spanning ± 1000 bp from enhancer centre, using either the standard MNase dose (**a**), or a four-fold excess (**b**). **c**, V-plot of Brg1 ChIP-seq. **d-e**, V-plots of histone H3 ChIP-seq in ES cells depleted of either Brg1 (**d**) or Chd4 (**e**), using shRNAs. **f-k**, V-plots of Oct4, Sox2, Nanog, TBP, H3K27ac and Pol II ChIP-seq. The standard MNase dose was used in all panels except in (**b**).



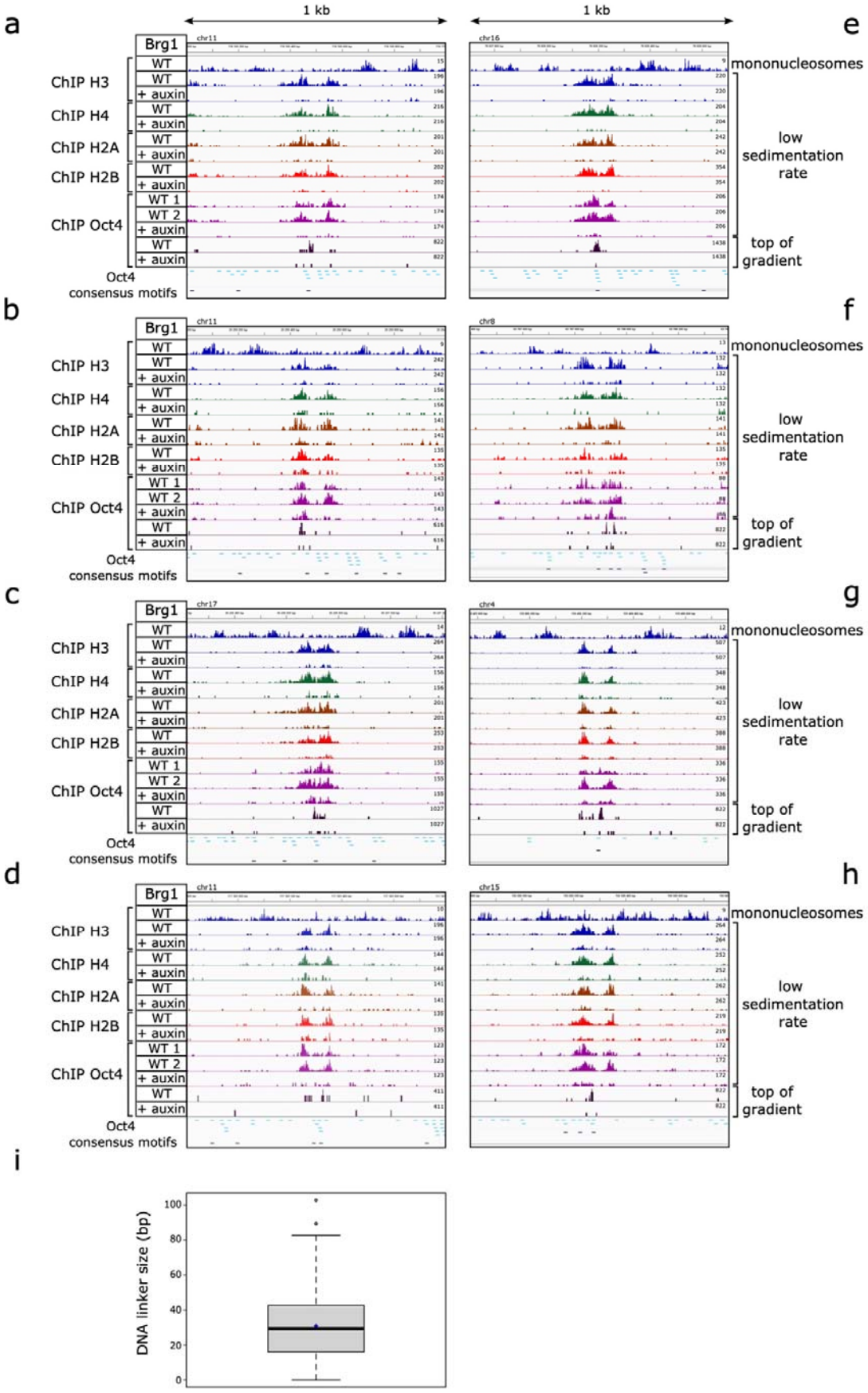
Extended Data Figure 3. Modular nucleosomal and subnucleosomal organization of enhancers in ES cells. Schematic representations of nucleosomes and subnucleosomal particles are shown below the V-plots of histone H3 ChIP-seq fragments spanning ± 1000 bp from the enhancer centre. **a**, Cluster 1 enhancers (left panel) are composed of a single basic module, defined by one fragile nucleosome at the centre and two subnucleosomal particles occupying the same genomic interval as the fragile nucleosome. Enhancers from cluster 2 (right panel) are composed of two basic modules, including two fragile nucleosomes in the central region between nucleosomes -1 and +1. These two fragile nucleosomes are separated by about 85 bp and thus cannot coexist on the same molecule. Each of the two fragile nucleosomes is associated with two subnucleosomal particles occupying the same genomic interval. Due to the proximity between the two modules, their

subnucleosomal particles partially overlap. (**b, c**) Enrichment for Oct4, Sox2, Nanog and Klf4 consensus binding sites, defined by centrimo software, within the same genomic intervals as in (**a**).

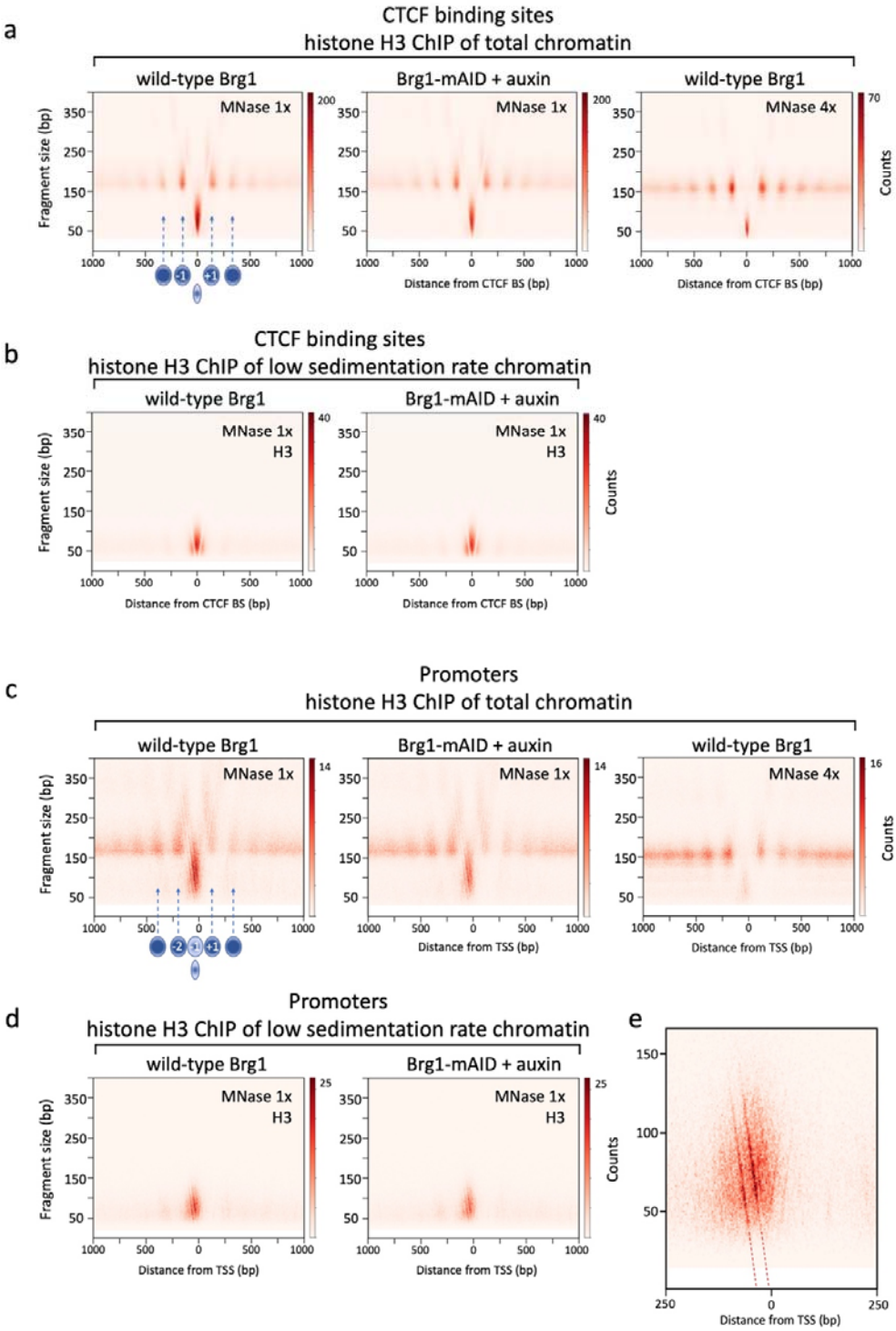


Extended Data Figure 4. Brg1 depletion induces a reduction in subnucleosomal particles at enhancers and a perturbation of nucleosome positioning. V-plots of ChIP-seq fragments spanning ± 1000 bp from the centre of cluster 1 enhancers. **a**, H3 and H2B ChIP-seq were performed on

MNase-digested chromatin prepared from ES cells expressing wild-type Brg1 (top panels) or after shRNA-mediated depletion of Brg1 (bottom panels). For the H3 ChIP-seq, depletion of Brg1 was obtained using an shRNA distinct from the shRNA used in Fig. 1. **b**, H3 ChIP-seq was performed using chromatin prepared from ES cells depleted of Brg1 using the AID system by either 3h or 20h of treatment with 1 mM auxin. **c**, Western blot analysis of protein extracts obtained from ES cells transfected with a plasmid expressing an shRNA targeting Brg1, from ES cells depleted of Brg1 using the AID system for either 20 or 3 h, or control cells. Ponceau staining of the membrane is shown as a loading control. The right panel shows the means ± sd of Brg1 protein levels in replicate experiments (n shRNA = 8, n 20h auxin = 3 and n 3h auxin = 4).

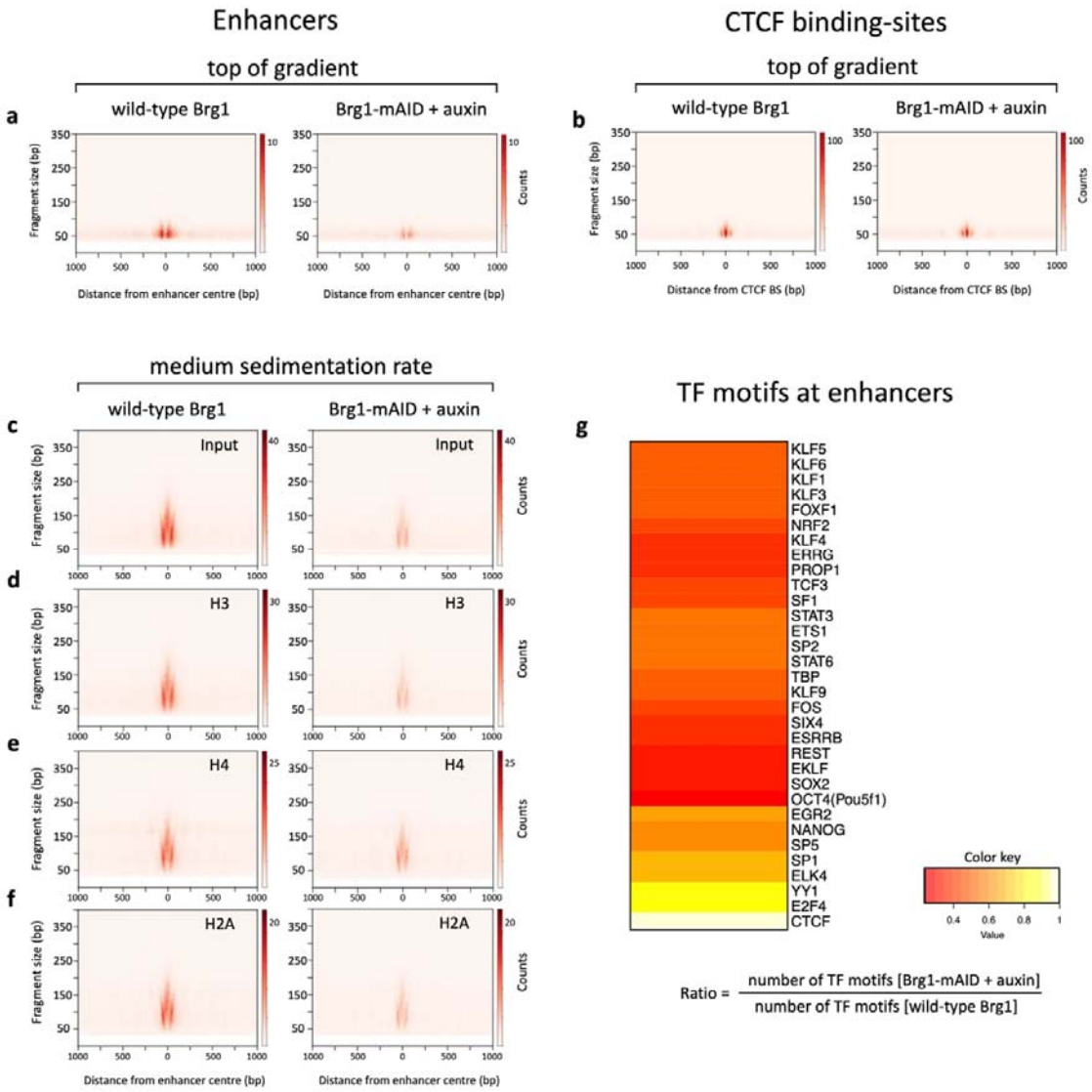

a
b
c
d
e
f
g
h
i
1 kb
1 kb
Brg1
ChIP H3
ChIP H4
ChIP H2A
ChIP H2B
ChIP Oct4
WT
+ auxin
WT 1
WT 2
Oct4
consensus motifs
mononucleosomes
low
sedimentation
rate
top of
gradient
DNA linker size (bp)
100
80
60
40
20
0

Extended Data Figure 5. Single-module enhancers accommodate one pair of 50-80 bp subnucleosomes. a-h, Density graphs showing the distribution of ChIP DNA fragment centres at representative examples of enhancers from clusters 1 and 8. The top lane of each panel shows the positions of canonical nucleosomes detected by histone H3 ChIP-seq of sucrose gradient fractions 11 and 13. The next eight lanes show the distribution of 50-80 bp subnucleosomes, revealed by histone H3, H4, H2A and H2B ChIP-seq of sucrose gradient fractions 5-6. The following five lanes show Oct4 enrichment detected by ChIP-seq of 50-80 bp subnucleosomes (fractions 5-6) or histone-free DNA from the top gradient fractions (fractions 2-3-4). Experiments were performed with chromatin prepared from ES cells expressing wild-type (WT) Brg1 and from cells depleted of Brg1 using the AID system (+ auxin). Two distinct antibodies against Oct4 were used in lanes WT 1 and 2. High ($p = 0.001$) and low ($p = 0.01$) stringency Oct4 consensus binding motifs are indicated in black and blue, respectively. **i**, Size distribution of the DNA linker that separates the two 50-80 bp subnucleosomes in single-module enhancers ($n = 496$, median = 29 bp). The box ranges from the first to the third quartile, the line across the box indicates the median, the blue diamond the mean, the whiskers show the maximum and minimum values of the distribution, and open circles represent outliers.



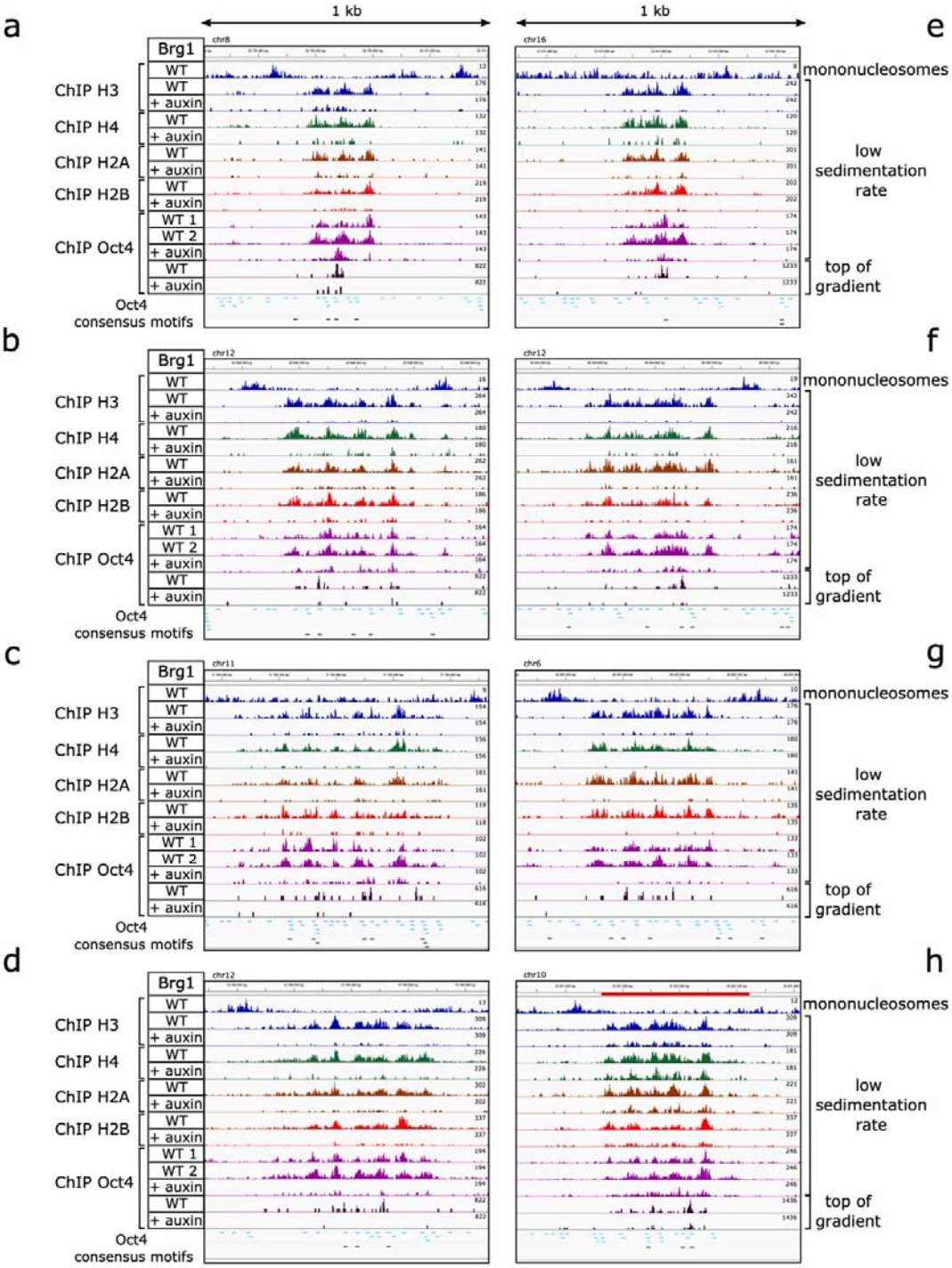
Extended Data Figure 6. Nucleosomal and subnucleosomal organization at promoters and CTCF binding sites. a, **c**, V-plots of histone H3 ChIP-seq fragments spanning ± 1000 bp from CTCF binding sites (**a**) and TSSs (**c**), using chromatin prepared from ES cells expressing wild-type Brg1 or

from cells depleted of Brg1 using the AID system. Chromatin digestion was performed with the standard MNase dose (left and mid panels) or a four-fold excess (right panel). **b**, **d**, MNase-fragmented chromatin, prepared from ES cells expressing wild-type Brg1, or from cells depleted of Brg1 using the AID system, was centrifugated through a sucrose gradient as in Fig. 2. V-plots of histone H3 ChIP-seq experiments performed with sucrose gradient fractions 5-6 reveal 50-80 bp subnucleosomes at CTCF binding sites (**b**) and promoters (**d**). The schematic illustrations in (**a**) and (**c**) indicate the positions of nucleosomes and subnucleosomal particles. **e**, Enlargement of **d** (left panel). The two left diagonals result from fragments cleaved precisely on the left and right sides of a region protected from MNase digestion. Extrapolation of the diagonals to y = 0 identifies this region as the -35 to -5 bp interval relative to the TSS. Protection might be conferred by low molecular weight components of the preinitiation complex.



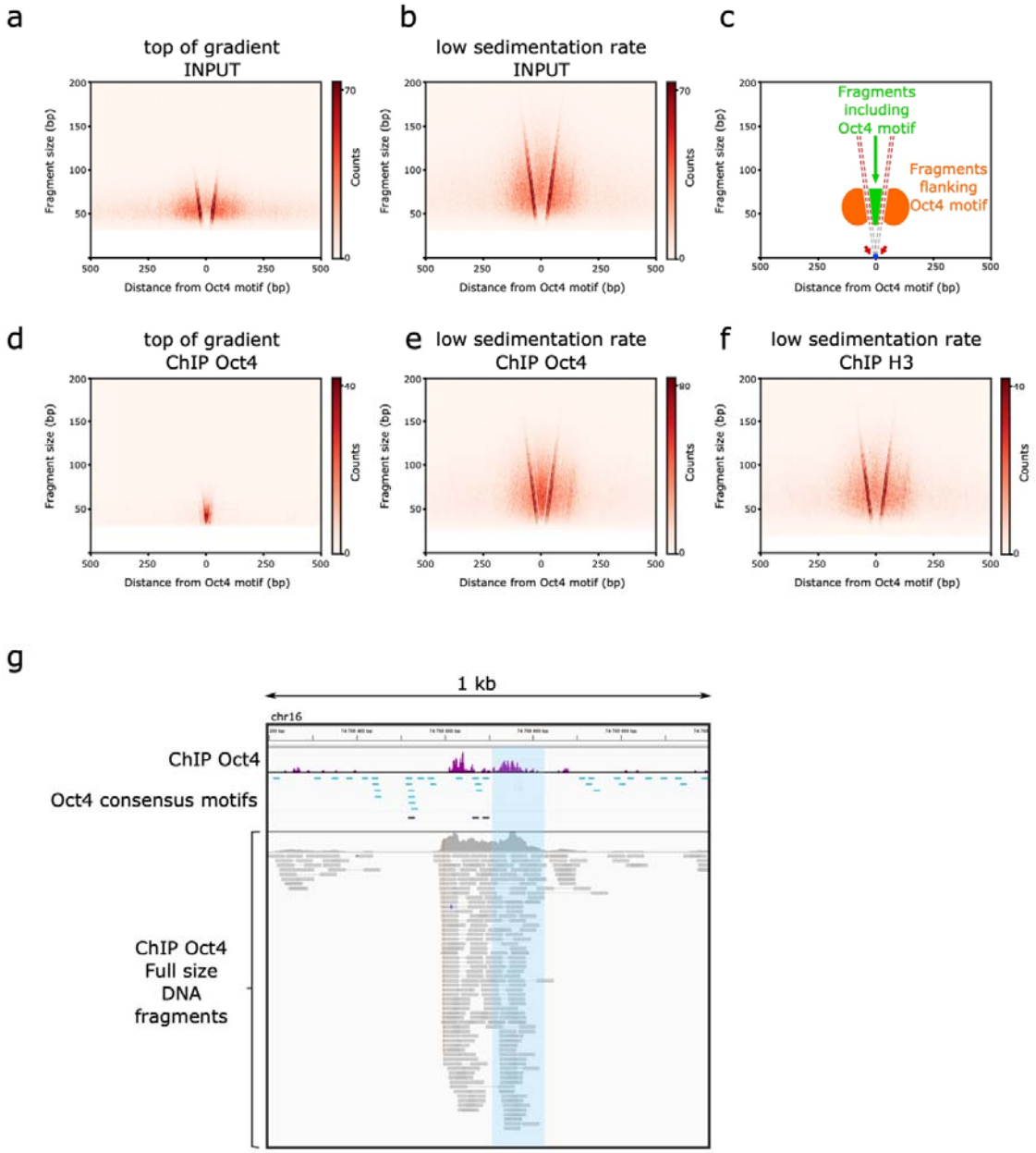
Extended Data Figure 7. Brg1 depletion alters the production of medium sedimentation rate subnucleosomal particles and accessible DNA at enhancers. Chromatin, prepared from ES cells expressing wild-type Brg1, or from cells depleted of Brg1 using the AID system, was centrifugated through a sucrose gradient as in Fig. 2. **a**, **b** V-plots of DNA fragments purified from the top gradient fractions spanning ± 1000 bp from the enhancer cluster 1 centre (**a**), or CTCF binding sites (**b**). **c**, V-plots of DNA fragments purified from medium sedimentation rate sucrose gradient fractions (pool of fractions 7-8-9) spanning ± 1000 bp from enhancer cluster 1 centre. **d-f**, V-plots of histone H3, H4 and H2A ChIP-seq experiments performed with the pool of medium sedimentation rate fractions indicated in (**c**) as input. The colour scales of each pair of panels were adjusted according to the total number of DNA fragments obtained in each dataset. **g**, TF motif detection with the HOMER tool on the DNA fragments revealed in (**a**). The ratio between the number of TF motifs detected in Brg1-depleted

versus Brg1-expressing samples was calculated and visualized as a heatmap. The number of CTCF motifs, which was invariant, was used for normalization.



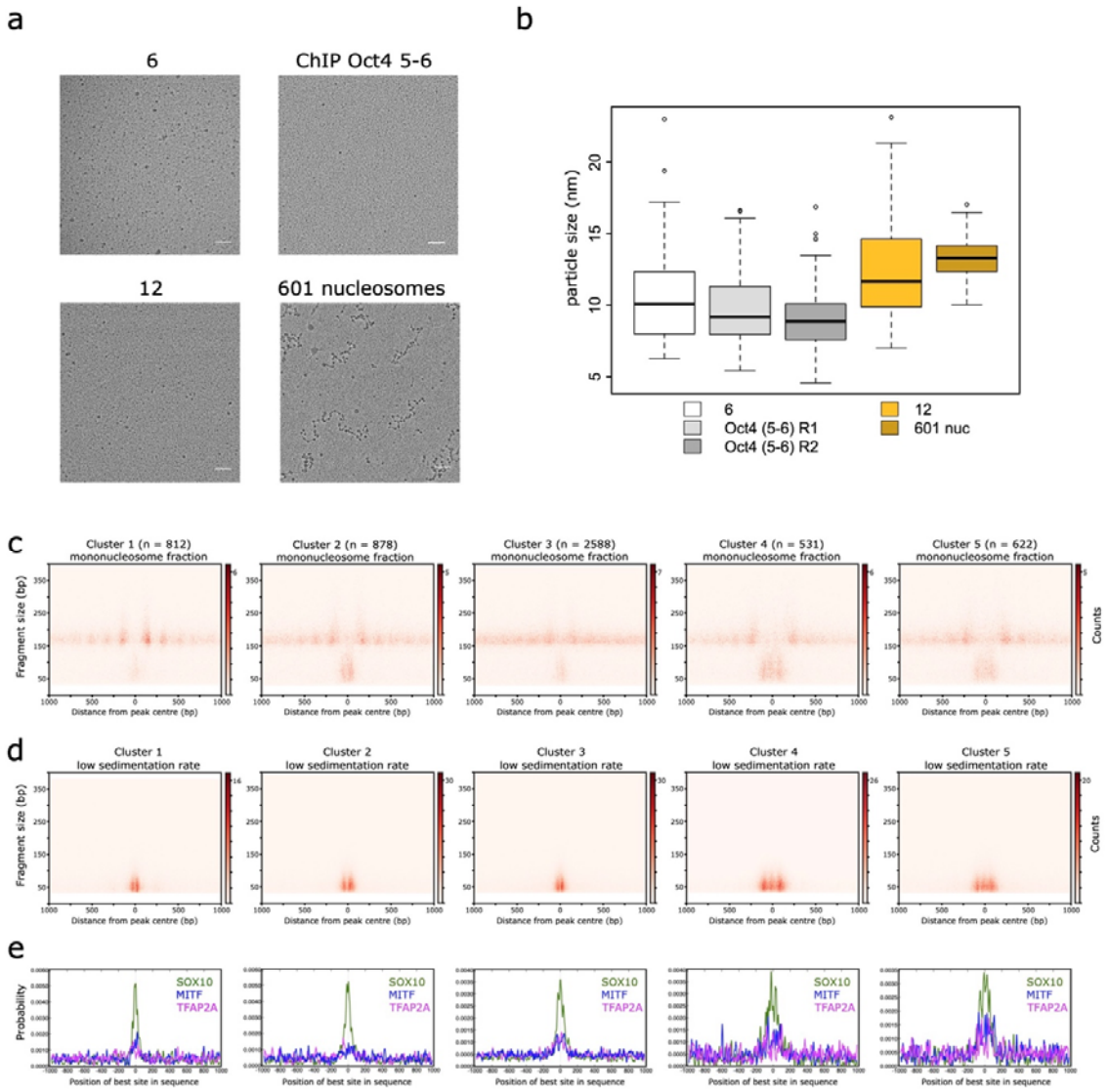
Extended Data Figure 8. Oct4 enrichment at enhancers coincides with the position of 50-80 bp subnucleosomes. a-g, Density graphs showing the distribution of ChIP DNA fragment centres at examples of multi-module enhancers from clusters 2 and 3. **h**, Example of an enhancer present within a super-enhancer, showing an organisation similar to multi-module enhancers. The top lane of each panel shows the positions of canonical nucleosomes detected by histone H3 ChIP-seq of sucrose

gradient fractions 11 and 13. The next eight lanes show the distribution of 50-80 bp subnucleosomes, revealed by histone H3, H4, H2A and H2B ChIP-seq of sucrose gradient fractions 5-6. The following five lanes show Oct4 enrichment detected by ChIP-seq of 50-80 bp subnucleosomes (fractions 5-6) or histone-free DNA from the top gradient fractions (fractions 2-3-4). Experiments were performed with chromatin prepared from ES cells expressing wild-type (WT) Brg1 and from cells depleted of Brg1 using the AID system (+ auxin). Two distinct antibodies against Oct4 were used in lanes WT 1 and 2. High ($p = 0.001$) and low ($p = 0.01$) stringency Oct4 consensus binding motifs are indicated in black and blue, respectively.



Extended Data Figure 9. Oct4 interacts with 50-80 bp subnucleosomes lacking its consensus binding motif. a-b and **d-f**, V-plots of enhancers centred on their Oct4 motif. Enhancers containing a single Oct4 consensus binding motif were selected for this analysis (n = 1,302). MNase-fragmented chromatin, prepared from ES cells expressing wild-type Brg1, was centrifugated through a sucrose gradient as in Fig. 2. **a-b**, V-plots of DNA fragments purified from the top gradient (**a**) and low sedimentation rate (**b**) fractions. **c**, Schematic diagram illustrating key features of V-plots centred on Oct4 motifs. The green area is defined by MNase-released DNA fragments containing the central Oct4 consensus binding motif. The two orange areas contain adjacent or proximal DNA fragments that do not contain the Oct4 motif. The red arrows indicate MNase hypersensitive sites on each side of the Oct4 motif, which are at the origin of the diagonals forming the 'V' on each map. **d-e**, Maps of Oct4

ChIP-seq experiments of the top gradient (**d**) and low sedimentation rate gradient (**e**) fractions, using as input the fractions shown in (**a**, **b**). **f**, Map of histone H3 ChIP-seq experiment of low sedimentation rate gradient fractions. **g**, The top lane shows the density graph of Oct4 ChIP-seq DNA fragments centres from 50-80 bp subnucleosomes at an example of a single-module enhancer. High ($p = 0.001$) and low ($p = 0.01$) stringency Oct4 consensus binding motifs are indicated in black and blue, respectively. Below is shown the density graph of Oct4 ChIP-seq full-size DNA fragments and the genomic position of each individually sequenced fragment immunoprecipitated with the Oct4 antibody. Blue shading highlights the DNA fragments that do not contain Oct4 motifs.



Extended Data Figure 10. EM analysis and conservation in human cells of 50-80 bp subnucleosomes. a, EM analysis of ES cell 50-80 bp subnucleosomes. The top panels show 50-80 bp subnucleosomes from sucrose gradient fraction 6 and particles from a pool of fractions 5-6 immunopurified with Oct4 antibodies. The bottom panels display mononucleosomes from fraction 12 and nucleosomes assembled *in vitro* on 601 DNA arrays. **b**, Boxplot shows particle size distribution in the fractions indicated in (**a**). The line across each box indicates the median, boxes indicate the first and third quartile, whiskers the distribution's maximum and minimum values, and open circles outliers. Two independent Oct4 immunopurification experiments were analysed. **c-e,** Nucleosomal and subnucleosomal organization of enhancers is conserved in human cells. MNase-digested chromatin prepared from human melanoma 501 Mel cells was centrifugated through a 10-30% sucrose gradient, as in Fig. 2. Histone H3 ChIP-seq experiments were carried out with gradient fractions containing either mononucleosomes or 50-80 bp subnucleosomes. These ChIP-seq datasets were used to perform k-means clustering of putative enhancers based on nucleosome and subnucleosome distribution

patterns, resulting in the isolation of five clusters of enhancers. **c**, **d**, V-plots of histone H3 ChIP-seq experiments performed with gradient fractions containing mononucleosomes (**c**) or 50-80 bp subnucleosomes (**d**), spanning ± 1000 bp from the centre of putative enhancers from each of the five clusters. **e**, Enrichment for SOX10, MITF and TFAP2A consensus binding sites, defined by Centrimo software, within the same genomic intervals as in (**c**, **d**).

MAIN REFERENCES

1. Kadoch, C. *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet* **45**, 592–601 (2013).
2. Iurlaro, M. *et al.* Mammalian SWI/SNF continuously restores local accessibility to chromatin. *Nat Genet* **53**, 279–287 (2021).
3. Schick, S. *et al.* Acute BAF perturbation causes immediate changes in chromatin accessibility. *Nat Genet* **53**, 269–278 (2021).
4. Chen, X. *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106–1117 (2008).
5. Sönmezer, C. *et al.* Molecular Co-occupancy Identifies Transcription Factor Binding Cooperativity In Vivo. *Molecular Cell* **81**, 255-267.e6 (2021).
6. Singh, G. *et al.* A flexible repertoire of transcription factor binding sites and a diversity threshold determines enhancer activity in embryonic stem cells. *Genome Res.* **31**, 564–575 (2021).
7. King, D. M. *et al.* Synthetic and genomic regulatory elements reveal aspects of cis-regulatory grammar in mouse embryonic stem cells. *eLife* **9**, e41279 (2020).
8. Spitz, F. & Furlong, E. E. M. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* **13**, 613–626 (2012).
9. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
10. Koch, F. *et al.* Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat Struct Mol Biol* **18**, 956–963 (2011).
11. Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21931–21936 (2010).
12. Yue, F. *et al.* A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**, 355–364 (2014).
13. Carone, B. R. *et al.* High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. *Dev Cell* **30**, 11–22 (2014).
14. Ishii, H., Kadonaga, J. T. & Ren, B. MPE-seq, a new method for the genome-wide analysis of chromatin structure. *Proc Natl Acad Sci U S A* **112**, E3457-3465 (2015).
15. Iwafuchi-Doi, M. *et al.* The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. *Molecular Cell* **62**, 79–91 (2016).
16. Lai, B. *et al.* Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. *Nature* **562**, 281–285 (2018).
17. Schones, D. E. *et al.* Dynamic Regulation of Nucleosome Positioning in the Human Genome. *Cell* **132**, 887–898 (2008).
18. Teif, V. B. *et al.* Genome-wide nucleosome positioning during embryonic stem cell development. *Nat Struct Mol Biol* **19**, 1185–1192 (2012).
19. Valouev, A. *et al.* Determinants of nucleosome organization in primary human cells. *Nature* **474**, 516–520 (2011).
20. Voong, L. N. *et al.* Insights into Nucleosome Organization in Mouse Embryonic Stem Cells through Chemical Mapping. *Cell* **167**, 1555-1570.e15 (2016).
21. Henikoff, J. G., Belsky, J. A., Krassovsky, K., MacAlpine, D. M. & Henikoff, S. Epigenome characterization at single base-pair resolution. *Proc Natl Acad Sci U S A* **108**, 18318–18323 (2011).
22. Clapier, C. R., Iwasa, J., Cairns, B. R. & Peterson, C. L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* **18**, 407–422 (2017).
23. de Dieuleveult, M. *et al.* Genome-wide nucleosome specificity and function of

chromatin remodellers in ES cells. *Nature* **530**, 113–116 (2016).
24. Markert, J. & Luger, K. Nucleosomes Meet Their Remodeler Match. *Trends Biochem Sci* **46**, 41–50 (2021).
25. Narlikar, G. J., Sundaramoorthy, R. & Owen-Hughes, T. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* **154**, 490–503 (2013).
26. Wang, W. *et al.* Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev* **10**, 2117–2130 (1996).
27. King, H. W. & Klose, R. J. The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. *Elife* **6**, (2017).
28. Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M. T. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep* **15**, 210–218 (2016).
29. Ramachandran, S., Ahmad, K. & Henikoff, S. Transcription and Remodeling Produce Asymmetrically Unwrapped Nucleosomal Intermediates. *Mol Cell* **68**, 1038-1053.e4 (2017).
30. Rhee, H. S., Bataille, A. R., Zhang, L. & Pugh, B. F. Subnucleosomal Structures and Nucleosome Asymmetry across a Genome. *Cell* **159**, 1377–1388 (2014).
31. Zlatanova, J., Bishop, T. C., Victor, J.-M., Jackson, V. & van Holde, K. The nucleosome family: dynamic and growing. *Structure* **17**, 160–171 (2009).
32. Fatmaoui, F. *et al. Cryo-electron tomography and deep learning denoising reveal native chromatin landscapes of interphase nuclei*. http://biorxiv.org/lookup/doi/10.1101/2022.08.16.502515 (2022) doi:10.1101/2022.08.16.502515.
33. Fan, H.-Y., He, X., Kingston, R. E. & Narlikar, G. J. Distinct strategies to make nucleosomal DNA accessible. *Mol Cell* **11**, 1311–1322 (2003).
34. Noll, H. & Noll, M. Sucrose gradient techniques and applications to nucleosome structure. *Methods Enzymol* **170**, 55–116 (1989).
35. Clément, C. & Almouzni, G. MCM2 binding to histones H3-H4 and ASF1 supports a tetramer-to-dimer model for histone inheritance at the replication fork. *Nat Struct Mol Biol* **22**, 587–589 (2015).
36. Ramachandran, S. & Henikoff, S. Nucleosome dynamics during chromatin remodeling *in vivo*. *Nucleus* **7**, 20–26 (2016).
37. Farnung, L., Ochmann, M., Garg, G., Vos, S. M. & Cramer, P. Structure of a backtracked hexasomal intermediate of nucleosome transcription. *Molecular Cell* S1097276522006104 (2022) doi:10.1016/j.molcel.2022.06.027.
38. Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* **276**, 19–42 (1998).
39. Zhou, K., Gaullier, G. & Luger, K. Nucleosome structure and dynamics are coming of age. *Nat Struct Mol Biol* **26**, 3–13 (2019).

METHODS

Cell culture

Mouse ES cells (46C, E14Tg2a and their derivatives) were grown on mitomycin C-inactivated mouse embryonic fibroblasts at 37 °C, 5 % CO_2, in DMEM (Sigma) supplemented by leukaemia inhibitory factor (LIF), 1X non-essential amino acids (Invitrogen), 15 % foetal calf serum (Invitrogen), 0.2 % β-mercaptoethanol (Sigma) and 1X penicillin/streptomycin (Invitrogen)[40]. The human melanoma 501Mel cell line[41] was grown in RPMI 1640 medium (Sigma) supplemented with 10% foetal calf serum and 1X penicillin/streptomycin.

Plasmid construction

We designed the shRNA-expressing plasmids as previously described[42]. Sequences of the sense strand of shRNAs targeting Brg1 are as follows: shRNA O5: 5'-GCTCCAGTAAAGATGTCTACT-3'; shRNA O7: 5'-GAGCGAATGCGGAGGCTTATG-3'. Sequences targeting *Chd4* were already described[23].

The AID system is based on the fusion, at the C-terminus of the target protein, of a mini AID (mAID) tag (68 amino-acids, 7.4 kDa)[28]. We adopted the two-selection marker strategy to target the two *Smarca4* alleles with the sequence encoding mAID[28]. We assembled the *Smarca4*-mAID targeting vector by a serial modification of the base vectors pMK287 (mAID-Hygro, Addgene #72825) and pL452[23]. First, we introduced two adapters, upstream and downstream of the region coding for mAID and Hygromycin resistance into the pMK287 vector. The upstream adapter contained a sequence encoding a $(GGGGS)_3$ linker in front of the mAID coding sequence. The region containing the linker-mAID sequence was then amplified by PCR using Phusion polymerase (NEB) from the modified pMK287-mAID-Hygro plasmid and subcloned into the pL452 vector which harbours the Neomycin resistance cassette under the control of a PGK promoter. We amplified the left and right homology arms (400-500 bp each) corresponding to the last exon and 3'UTR of *Smarca4* by PCR from E14Tg2A genomic DNA. We designed the oligonucleotides that amplify homology arms to introduce silent mutations after integrating the mAID cassette downstream of the *Smarca4* gene, to prevent re-cutting by the Cas9 enzyme. In the next cloning steps, we introduced the PCR products of left and right homology arms using restriction enzyme digestions into both vectors (pMK287-mAID-Hygro and pL452-mAID-Neo).

We assembled the CRISPR/Cas9 vectors by annealing pairs of oligos carrying the sgRNA sequences and cloning them into pSpCas9(BB)-2A-Puro (PX459) (Addgene #62988) as described[43]. Two distinct *Smarca4*-targeting sgRNAs were cloned separately by annealing oligos caccgTTGGCTGGGACGAGCGCCTC and aaacGAGGCGCTCGTCCCAGCCAAc for the first sgRNA and caccgGCGCCTCGGGGTCAGGACTC and aaacGAGTCCTGACCCCGAGGCGCc for the second sgRNA. We designed the sgRNAs to target the area close to the last exon of *Smarca4* using the Zhang lab guide design tool (https://zlab.bio/guide-design-resources).

Generation of the Smarca4-mAID cell line

To generate stable cell lines, we used the E14Tg2a-Tir1 cell line, which was obtained by targeting the *Tigre* locus of E14Tg2a cells with a transgene expressing *Os* Tir1[44]. Using the Neon electroporation kit (ThermoFisher), we transfected this cell line with the two vectors encoding the Cas9 enzyme and sgRNAs (2.5 µg each), and with the two linearized plasmids carrying the *Smarca4* homology arms and the mAID tag (pMK287, pL452, 10 µg each). Neomycin (250 µg/ml, G-418, Sigma) and Hygromycin B (300 µg/ml, Thermo Fisher Scientific) selection was started after 24 h. After six days of selection, single clones were manually picked, expanded in 24-well plates and subjected to genotyping by PCR and sequencing of the knock-in region. Western blotting of protein extracts from auxin-treated cells confirmed selected clones with a homozygous insertion of the mAID cassette at the *Smarca4* locus.

Chromatin remodeller depletion

For shRNA-mediated depletion, we transfected two million ES cells by electroporation using Amaxa nucleofector (Lonza) or Neon kits, with 15 µg of pHYPER-linker control plasmid or a pHYPER plasmid containing remodeller-specific shRNA. Puromycin (2 µg/ml) was added to the culture medium 24h after electroporation, and the cells were grown under selection for an additional 30h (Chd4 shRNA) or 48h (Brg1 shRNA) period, as previously described[23]. Western blot experiments validated the efficiency of remodeller depletion.

For auxin-mediated protein depletion, we supplemented the cell culture medium of cells expressing Brg1-mAID with 1 mM auxin (IAA, indole-3-acetic acid, Cayman Chemical) for 3 h or 20 h. All auxin-mediated depletion of Brg1 experiments described in the manuscript were realized at the 20h auxin timepoint unless specified otherwise.

Chromatin preparation

ES cells were treated with trypsin to obtain a single-cell suspension and separated from mouse embryonic fibroblast feeders by a 30 min incubation at 37°C in D15 medium on a gelatin-coated culture plate. ES cells were collected and fixed in D15 culture medium containing 1 % formaldehyde (Sigma) for 10 min at 20°C. 501Mel cells were treated with trypsin and fixed in culture medium as described for ES cells. We stopped fixation by adding glycine to 125 mM, and the cells were centrifugated and washed three times with PBS. Cells were then permeabilized in Tris-HCl pH 7.5 15 mM, NaCl 15 mM, $MgCl_2$ 5 mM, EGTA 0.1 mM, KCl 60 mM, Sucrose 0.3 M, 0.4% IGEPAL CA-630 (I3021, Sigma) during 15 min on ice. We next diluted the cells with two volumes of MNase buffer (Tris-HCl pH 7.5 20 mM, NaCl 20 mM, $CaCl_2$ 2 mM, $MgCl_2$ 4 mM, KCl 15 mM, 0.1 M sucrose), and chromatin was digested during 10 min at 37 °C with 6 Kunitz units of MNase (New England Biolabs, 200 Kunitz units/µl) per 1 million cells. We stopped MNase digestion by adding a final concentration of 10 mM EDTA. We released chromatin from MNase-treated ES cells by passing the cell suspension 13 times through a 26G syringe. Next, we centrifuged the samples, and the supernatants containing the solubilized chromatin were used directly for ChIP-seq experiments or sucrose gradient centrifugation. The quality and reproducibility of the MNase digestion pattern of all samples were validated after crosslink reversal by electrophoresis on a 1.7 % agarose gel.

Fractionation of MNase-digested chromatin by sucrose gradient centrifugation

Chromatin samples prepared from 40 million ES or 501 Mel cells were supplemented with protease inhibitors (Complete Protease Inhibitor Cocktail, Roche) and loaded on a 9.9 ml manually poured 10-30 % sucrose gradient containing 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 mM EGTA, 80 mM NaCl and protease inhibitors. We centrifuged the chromatin samples for 16 h at 4 °C, 40,000 rpm (197,000 g) using a Beckman Coulter SW 41 swinging-bucket rotor. After centrifugation, 20 chromatin fractions of 500 µl were collected from top to bottom of the tube and stored at -80°C. Aliquots of each chromatin fraction were treated with RNAse A and proteinase K, and subjected to reverse crosslinking for 16 h at 65 °C, followed by phenol-chloroform DNA extraction and ethanol precipitation. DNA was quantified using the Qubit system (Thermo) and analysed on a 4% agarose gel or used for library preparation and deep sequencing.

MNase ChIP-seq

We realized all ChIP-seq experiments with formaldehyde-fixed, MNase-fragmented chromatin, either with individual sucrose gradient fractions, pools of gradient fractions, or whole-cell chromatin extracts. For the latter, we used chromatin solubilized from 1.2 million ES cells for ChIP-seq with antibodies against histones or from 10 million cells for ChIP-seq with antibodies against TFs and Pol II. Chromatin samples were incubated with antibodies (20 µg of antibodies against TFs or Pol II and 5 µg of antibodies against histones) overnight at 4 °C. Chromatin-antibody complexes were next incubated with 50 µl of protein G-agarose beads for 4 h at 4 °C. Then, chromatin-antibody-beads complexes were washed three times with TEN buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 0.01% Igepal), four times with WBLiCl buffer (50 mM Hepes pH 7.5, 500 mM LiCl, 1 mM EDTA, 1 % IGEPAL, 0.7 % Na-deoxycholate) and once with TEN buffer. Between each wash, we centrifuged the beads for 3 min at 800 g, 4 °C. Elution was realized by resuspending the beads in elution buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 1% SDS). For ChIP-seq experiments realized with the antibody targeting the 20 C-terminal amino acids of Oct4 (Fig. 4c and WT 1 lanes of Fig. 4d, Extended Data Fig. 5 and 8), we performed the washes and elution by peptide competition as described below for sequential ChIP-seq of Oct4- and H2B-bound chromatin. Eluted chromatin was treated with proteinase K and subjected to reverse crosslinking and phenol/chloroform extraction, followed by ethanol precipitation. We prepared DNA libraries using MicroPlex Library Preparation Kit v2 or v3 (Diagenode). Sequencing was performed onto a NextSeq 500 Illumina platform using NextSeq 500 High Output v2 (75 cycles) kit. At least two biological replicates were realized and sequenced for each histone or TF ChIP-seq experiment.

Sequential ChIP-seq of Oct4- and H2B-bound genomic DNA

Chromatin prepared from 240 million E14Tg2a ES cells was used for each experiment. We incubated the chromatin with either 45 µg of antibodies against Oct4 (MBS420786, MyBioSource) or with 45 µg of non-specific IgG as a control. After 16 h at 4 °C, we added Protein G-agarose beads to chromatin-antibody complexes and further incubated for 4 h at 4 °C. After seven washes with TEN buffer, we incubated the beads with TEN buffer containing Oct4 peptide (C-terminal part of Oct4 sequence: H-Cys-Lys-Lys-Lys-Lys-Pro-Ser-Val-Pro-Val-Thr-Ala-Leu-Gly-Ser-Pro-Met-His-Ser-Asn-OH) for elution by competition. We performed three consecutive elution steps: 2 h at room temperature with TEN buffer containing 2 mg/ml of Oct4 peptide, 2 h at room temperature with 1 mg/ml of peptide, and 16

h at 4 °C with 1 mg/ml peptide. The three elution fractions were pooled and incubated with 5 µg of antibodies against histone H2B for 16 h at 4 °C. Protein G-agarose beads were next added and incubated for 4 h at 4 °C. We performed the washes and elution steps of this second ChIP as described above for regular ChIP-seq experiments. Two biological replicates were obtained for each Oct4 and non-specific IgG antibody.

Transmission Electronic Microscopy

Chromatin samples were prepared from 80 million E14Tg2a ES cells and centrifugated through a sucrose gradient, as described above for MNase ChIP-seq. We examined subnucleosomal particles collected from fractions 5, 6, 11 and 12 of the sucrose gradient and particles further purified from a pool of fractions 5-6 by immunoprecipitation with the antibody targeting the 20 C-terminal amino-acids of Oct4, as described above. We analysed half of each replicate sample of Oct4-immunopurified particles by EM as shown in ED Fig. 10a, b, and we processed the other half as a ChIP-seq sample, which was used for the V-plot shown in Fig. 4c, and for the WT1 lane of the density graphs shown in Fig. 4d and Extended Data Fig. 5 and 8.

Aliquots of sucrose gradient fractions were dialyzed overnight into TEAP10 buffer (10 mM Triethanolamine-HCl (pH 7.5), 0.2 mM EDTA, 10 mM NaCl, 0.2 mM PMSF), then crosslinked with 0.1% glutaraldehyde at 4°C for 2 h and quenched by the addition of Tris (pH 8) to a final concentration of 50 mM. Samples were coated with 2 x 10^{-4} % benzalkonium chloride (BAC) for 1 h at room temperature, and 5 µl of the sample (approximately 0.02 ng/ µl) was adsorbed onto carbon/formvar supported 200 mesh copper grids (TAAB technologies). Grids were washed twice with deionized water, dehydrated with 90% ethanol, and blotted dry. Grids were rotary shadowed with 3 nm platinum at a 7° angle and examined on the JEOL JEM-1400Plus Transmission Electron Microscope.

We randomly selected the particles and measured them with Image J for statistical analysis of particle size. Most particles had a disk-like structure, and we measured them along the diameter. We measured particles having an elliptical shape along the major axis, as this likely represents the front view of the nucleosomal particles. We built the boxplot shown in Extended data Fig. 10b with the sizes of 92, 102, 97, 71 and 100 particles from one sample of sucrose gradient fraction 6, two independent samples of Oct4-immunopurified particles (Oct4 (5-6) R1 and R2), one sample of fraction 12 (mononucleosomes) and one preparation of *in vitro*-assembled 601 nucleosomes, respectively. To choose the appropriate statistical test to compare the different groups of particles, we first verified the size distribution in each group

and performed Shapiro-Wilk normality tests. This analysis revealed that the data collected for all groups of subnucleosomal particles did not display a normal distribution. We thus selected a Mann-Whitney test (R Wilcox.test function; two.sided; paired = FALSE; conf.level = 0.95 parameters) to determine if the sizes of subnucleosomal particles and canonical nucleosomes were significantly different. This statistical analysis revealed that subnucleosomal particles of each group are significantly different in size from canonical nucleosomes; particles from fraction 6 versus those from fraction 12: p-value = 6.6×10^{-05}; particles from Oct4 (5-6) R1 versus those from fraction 12: p-value = 1.5×10^{-07}; particles from Oct4 (5-6) R2 versus those from fraction 12: p-value = 2.9×10^{-11}, particles from fraction 6 versus 601 nucleosomes: p-value = 2.4×10^{-14}; particles from Oct4 (5-6) R1 versus 601 nucleosomes: p-value < 2.2×10^{-16}; particles from Oct4 (5-6) R2 versus 601 nucleosomes: p-value < 2.2×10^{-16}.

Construction of the hemisome-link-hemisome model

To generate the hemisome-link-hemisome molecular model, we separated the best resolution nucleosome structure (1kx5 bp) into two hemisomes by cutting the DNA at the dyad position. We positioned a B-DNA linker at the generated DNA extremity by overlapping four base pairs. A representative length of 29 bp was chosen, corresponding to the median distance between two hemisome-like particles at enhancers. The second hemisome was positioned similarly at the second extremity of the DNA linker. The exposed DNA is highly sensitive to MNase and thus digested prior to the ChIP experiment. The resulting isolated particles, corresponding to hemisomes, would contain 50-70 bp DNA fragments linked to histones H3-H4-H2A and H2B.

Computational analyses:

Lists of ES cell enhancers

We first mapped the distribution of cis-regulatory elements onto the ES cell genome using DNaseI-seq data (GSM1014187), resulting in the identification of 139,454 DNase-hypersensitive (DHS) regions, as previously described[23]. Gene promoters were filtered out using RefSeq coordinates of TSSs, and putative enhancers were identified based on the co-occurring binding of pluripotency-associated TFs Oct4 (GSM1082340), Sox2 (GSM1082341) and Nanog (GSM1082342), as well as of the mediator (GSM560347)[4,45]. To avoid potential interferences with CTCF-bound regions in our nucleosomal organization analysis, we removed from these putative enhancers those bound by CTCF, using ES cell datasets GSM723015 and GSM1828650. 19,365 loci displaying high ChIP-seq signal for Oct4, Sox2,

Nanog, and the Mediator, but negative for CTCF, were identified as putative enhancer elements. We then used the histone H3 MNase ChIP-seq datasets generated in this study to analyse nucleosomal and subnucleosomal organization at these putative enhancers. One of these H3 MNase ChIP-seq datasets was converted into a BED file and further subdivided into five files containing DNA fragments of 30-80, 81-110, 111-140, 141-180, and >181 bp. The number of DNA fragments in each subgroup was as follows. 30-80 bp: 13,806,227 DNA fragments; 81-110 bp: 9,480,961 fragments; 111-140 bp: 13,589,276 fragments; 141-180 bp: 64,258,774 fragments; >180 bp: 51,518,836 fragments. Next, in each BED file, DNA fragments were reduced in length to 10 bp, centred at the level of fragment midpoint. Using seqMINER[46], fragment densities from these modified BED files were collected in a -1500/+1500 bp window around the centre of each enhancer (defined in this study as the centre of the DHS peak) and subjected to several rounds of k-means clustering. This analysis allowed us to classify 10,304 (out of the initial 19,365) putative enhancers into the 12 clusters shown in Extended Data Fig. 1.

We developed the exPATT.R tool to detect Oct4 DNA binding motifs and change the genomic coordinates of each enhancer to have the Oct4 motif at its centre. The exPATT.R R script filters genomic regions, recorded in an input BED file, that match a given DNA pattern (Oct4 consensus motif in this study) depicted as a position weight matrix (PWM). These pattern matrices were matched in genomic region using the PWM-related functions of the Bioconductor package Biostrings[47] with a minimum score of 80% of the highest possible score. Then, genomic regions with at least one pattern hit were filtered. The output of this script was a new BEDGRAPH file with filtered DNA regions of 2 kb centred around the pattern. All BED file input and output operations were performed using the rtracklayer package. Then, we used bedtools to select enhancers bearing a single central Oct4 motif per enhancer (n = 1,302, Extended Data Fig. 9).

Lists of 501 Mel cell enhancers

Our extensive analysis of the distribution of 50-80 bp subnucleosomes in mouse ES cells revealed their accumulation at enhancers, promoters and CTCF-binding sites (this study). Based on the hypothesis that subnucleosomal particles are conserved in human cells, we mapped cis-regulatory elements in human 501 Mel cells using the histone H3 ChIP-seq dataset obtained from 50-80 bp subnucleosomes purified from these cells. We used MACS2 to call 91,205 peaks from this dataset. We next focused our analysis on putative enhancer elements located in intergenic regions. We filtered out all regions of the human genome

located within 3 kb of gene TSSs or 10 kb of gene end coordinates. We also set up a strategy to remove CTCF-bound loci from this list of peaks. Since the CTCF ChIP-seq distribution has not been defined in 501 Mel cells, we took advantage of the observation that CTCF binding to mammalian genomes is relatively independent of cell type[48,49]. We removed from the putative enhancer list all loci bound by CTCF in two distinct cancer cell lines, using ENCODE CTCF ChIP-seq datasets GSM749683, GSM749715, GSM749679 and GSM749768. This combination of filtering steps led to a list of 14,502 putative 501 Mel cell enhancers. K-means clustering of these 14,502 elements using the H3 ChIP-seq datasets from sucrose gradient-purified 50-80 bp subnucleosomes and from the fraction containing canonical nucleosomes allowed us to isolate the five clusters of putative 501 Mel cell enhancers shown in Extended data Fig. 10c-e.

List of CTCF binding sites

ES cell genomic CTCF binding sites were identified based on the ChIP-seq signal of several CTCF datasets (GSM723015, GSM1828650, GSM288351 and GSM560352) at DNase-hypersensitive (DHS) regions, as described above for enhancers. All loci proximal (< 2 kb) to gene promoters and putative enhancers were removed. We used a subset (n = 9,850) of these CTCF-positive regions to generate the V-plots shown in Extended data Fig. 6 and 7.

List of promoters

We rank-ordered gene promoters transcribed in mouse ES cells according to the distance between the two divergent TSSs located on the plus strand and the minus strand, using 5' ends of Start-RNA reads (GSE43390), as described[50]. ES cell Start-RNA reads datasets were first aligned to mm9 using Bowtie 1.3.0. Only the reads mapping to a single position with a maximum of two mismatches were conserved. We derived the reference promoter list from Refseq annotation. We defined promoter windows as ± 1000 bases around the TSS position. We discarded windows with more than one TSS on the plus strand. We also removed the promoters associated with less than five overlapping Start-RNA reads. We calculated the coverage using a bin of one base. We selected the first six bases of the 5-prime reads as the signal to adjust the position of the TSS on the plus strand. We used a Python script to find the nearest local maximum upstream of the Refseq annotation, which we defined as the TSS position in this study. Next, we determined the location of the antisense Start-RNA signal on the minus strand in the region located upstream of the plus strand TSS. As above, we selected the first six bases of the 5-prime reads to define the position of the antisense TSS. We then

designated the coordinate of the antisense TSS as the location having maximum Start-RNA expression on the minus strand, upstream of the sense TSS. We next sorted the promoters according to the distance between sense and antisense TSS coordinates. In this study, we focused our analysis on the 1,000 TSSs with the smallest distance between sense and antisense TSSs. The nucleosomal and subnucleosomal organization of the other groups of TSSs is similar and will be presented in detail in a separate publication.

MNase ChIP-seq and MNase-seq analysis

Quality controls, mapping and post-mapping process

The processes described below were done using Snakemake[51], Conda and in-house python scripts. We controlled the quality of sequencing data with FastQC (fastqc=0.11.9, default parameters)[52] and fastqScreen (fastq-screen=0.14.0) using bowtie2 as default aligner[53]. Depending on their quality, sequencing files were trimmed with fastp (fastp=0.20.1), using the process described hereafter. Reads with a size below 25 bp were discarded to keep mapping specificity. Adapter sequences were auto-detected over ~1 million reads and removed. Reads with a base quality under 10 through a window size of 6 bases were discarded, and poly-G sequences were filtered out. Potential artefactual enrichments were assessed over 1/20th of the raw reads and removed[54]. Trimmed paired-end reads were mapped with Bowtie 2 (bowtie2=2.4.2)[55] over the Gencode's 9th version of the mouse genome for ES cells or the 38th version of the human genome for 501 Mel cells. After mapping, reads were sorted by genomic coordinates using Samtools (samtools=1.11). Artefactual fragment duplication (optical and local) was assessed with Picard (picard=2.23.8; 'Picard Toolkit.' 2019. Broad Institute, GitHub Repository. http://broadinstitute.github.io/picard/; Broad Institute). Samtools was used to remove marked duplicates, unmapped, and reads involving a mapping quality below 30. The RPKM normalized genomic coverage was calculated with Deeptools for MNase data with a bin size of one base and centred on the fragment centre for a fragment size of 10 to 500 bp. Chromosomes M, Y and X were not used for the normalization. The effective mouse genome size was set to 2,652,783,500 bp. (deeptools=3.5.0)[56].

IGV visualizations and 2D histograms (V-plots).

Peaks and normalized coverage profiles were visualized as IGV (Integrative Genomics Viewer) screen-shots, heatmaps, and 2D graphs (V-plots). We converted BAM files containing paired-end mapped reads to BED files containing DNA fragments using Bedtools

bamtobed (bedtools=2.29.2)[57]. These DNA fragments were visualized as a heatmap with seqMINER software and as a 2D graph (V-plots) using a homemade script. In the plot, the y-axis corresponds to DNA fragment size in bp and the x-axis to the relative position of DNA centre fragments on the genome. Each dot corresponds to the coordinates of a DNA fragment midpoint, and colour intensity reflects the number of fragments at that position. When comparing several samples, we chose the colour intensity manually, and a ratio was applied to reduce the number of fragments to the same level in each sample.

TF motif detection in DNA fragments purified from the top gradient fractions.

We used MACS2 callpeak (macs2 2.2.7.1) to identify the genomic regions enriched in histone-free DNA fragments sequenced from the top gradient fractions (DNA purified from pools of sucrose gradient 2-3-4). We set the minimal size for a peak to 40 bp. Then, motifs were detected with HOMER findMotifsGenome.pl for mouse genome (homer 4.11). For Extended Data Fig. 7h, we selected 32 TF motifs with a significant p-value ($p < 10^{-50}$) in the samples prepared from ES cells expressing wild-type Brg1 (control samples). We noticed that the number of CTCF motifs detected was unaffected by depletion of Brg1, in agreement with previous observations based on ATAC-seq experiments[58]. We thus choose to normalise the data according to the number of CTCF motifs detected in each sample. Using a homemade R script, the number of CTCF motifs found in control (wild-type Brg1) samples was divided by the number of CTCF motifs detected in Brg1-depleted samples. This coefficient was then applied to 31 other TF motifs to calculate the ratio between the number of motifs detected in chromatin samples prepared from Brg1-depleted and Brg1-expressing ES cells.

Detection of TF motifs at enhancers.

We performed motifs analysis with Meme (v5.3.0)[59]. Detection of Oct4 motifs was carried out with Fimo over selected lists of enhancers, using two distinct p-values (0.001 and 0.01) to allow different levels of stringency in the motif sequence matching. The Oct4 motif matrix (MA0142.1) came from the Jaspar database[60]. We used Centrimo to define Oct4, Sox2, Nanog and Klf4 motif enrichment profiles at ES cell enhancers of clusters 1 and 2 (Extended Data Fig. 3b,c). For Centrimo analysis of putative enhancers in human 501 Mel cells, we selected the DNA binding motifs of SOX10, MITF and TFAP2A, which were shown to be enriched at enhancers[41].

Analysis of the frequency at which subnucleosomal particles are organized in pairs at enhancers and determination of DNA linker length.

IGV snapshots of ES cell enhancers from clusters 1 and 8 (n = 1,430) were analysed individually and separated into three subgroups based on the enrichment pattern of histone H3, H4, H2A and H2B ChIP-seq peaks for 50-80 bp subnucleosomes: i) 496 enhancers displayed a single-module organisation with two well-separated peaks, as shown in Fig. 4d and Extended Data Fig. 5; ii) 881 enhancers exhibited three or more discernible histone peaks, and were classified as multi-module enhancers (Extended Data Fig. 8); iii) 53 enhancers (3.7 % of total) presented a single histone peak that could be interpreted either as a rare subcategory of enhancers bearing a single subnucleosomal particle, or as single module enhancers having a pair of two closely located histone peaks fused into a single peak. For the three subgroups, the ChIP-seq peaks of histones H3, H4, H2A and H2B were always overlapping each other in genomic position, confirming the systematic association of the four histones in these 50-80 bp subnucleosomes**.** We measured the distance D in bp between the centres of the two peaks of the subgroup of 496 single-module enhancers. Assuming that each of these two peaks marks the position of a hemisome protecting an average of 64 bp of DNA from MNase digestion, as proposed in Fig. 4f, we calculated the linker length between these two putative hemisomes as D minus 64 bp. We found that this DNA linker's length ranges from 0 to 102 bp (median 29 bp) (Extended Data Fig. 5i).

METHODS REFERENCES

40. Tessarollo, L. Manipulating mouse embryonic stem cells. *Methods Mol Biol* **158**, 47–63 (2001).
41. Laurette, P. *et al.* Transcription factor MITF and remodeller BRG1 define chromatin organisation at regulatory elements in melanoma cells. *eLife* **4**, e06857 (2015).
42. Berlivet, S., Hmitou, I., Picaud, H. & Gérard, M. Efficient Depletion of Essential Gene Products for Loss-of-Function Studies in Embryonic Stem Cells. *Methods Mol Biol* **1622**, 91–100 (2017).
43. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281–2308 (2013).
44. Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930-944.e22 (2017).
45. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
46. Ye, T. *et al.* seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res* **39**, e35 (2011).
47. Wasserman, W. W. & Sandelin, A. Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet* **5**, 276–287 (2004).

48. Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
49. Wang, H. *et al.* Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res* **22**, 1680–1688 (2012).
50. Scruggs, B. S. *et al.* Bidirectional Transcription Arises from Two Distinct Hubs of Transcription Factor Binding and Active Chromatin. *Molecular Cell* **58**, 1101–1112 (2015).
51. Mölder, F. *et al.* Sustainable data analysis with Snakemake. *F1000Res* **10**, 33 (2021).
52. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
53. Wingett, S. W. & Andrews, S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res* **7**, 1338 (2018).
54. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
55. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).
56. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165 (2016).
57. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
58. Barisic, D., Stadler, M. B., Iurlaro, M. & Schübeler, D. Mammalian ISWI and SWI/SNF selectively mediate binding of distinct transcription factors. *Nature* **569**, 136–140 (2019).
59. Bailey, T. L. *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**, W202-208 (2009).
60. Fornes, O. *et al.* JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **48**, D87–D92 (2020).

Data availability

We deposited high-throughput sequencing data at the Gene Expression Omnibus with accession GSE210780, GSE210444 and GSE209914. Density graphs used in Fig. 4 and Extended Data Fig. 5 and 8 have been deposited at the Zenodo Data repository: https://zenodo.org/record/7056534.

Acknowledgements:

We thank A. Goldar, A. Leforestier, J.M. Victor, S. Khochbin, K. Yen, S. Mahony, I. Davidson for discussions, E. Nora and B. Bruneau for the E14Tg2a-Tir1 cell line, and C. Mann for reading the manuscript.

This work was supported by grants from the INCA (2017-1-PL BIO-02-CEA-1), Investissement d'Avenir ANR (Revive Labex ANR-10- LABX-73), ANR (ANR-17-CE12-0001-01), PRC/PICS07434 and Fondation ARC pour la recherche sur le cancer www.fondation-arc.org.

Author's contributions:

M.C.N. designed, performed and analysed experiments and bioinformatic analyses.

A.M.K. designed, performed and analysed experiments.

E. D. achieved the majority of bioinformatic analyses.

W. A. performed and analysed EM experiments.

C.D. and F.R. performed experiments.

H.P. assisted with experiments.

O.A. performed bioinformatic analyses.

J.C.A. designed tools for bioinformatic analyses.

N.G. supervised EM experimentation and analysis

F.O. performed modelling of the split-nucleosome and contributed to data interpretation.

S.C. designed and analysed experiments.

M.G. conceived the project, designed and analysed experiments, and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information (containing supplementary information line)

Supplementary tables:

Supplementary Table 1: Antibodies used in this study

Supplementary Table 2: cell lines used in this study

Supplementary Table 3: plasmids used in this study

Supplementary Table 4: sequencing depth

Correspondence and requests for material should be addressed to M.G.