Aus dem Biomedizinischen Centrum
der Ludwig-Maximilians-Universität München,
Lehrstuhl Molekularbiologie
Vorstand: Prof. Dr. Peter B. Becker

Absolute nucleosome occupancy and reconstitution of nucleosome positioning mechanisms for *Saccharomyces cerevisiae*

Dissertation
zum Erwerb des Doktorgrades der Naturwissenschaften
an der Medizinischen Fakultät
der Ludwig-Maximilians-Universität zu München

vorgelegt von
Anna Elisa Barbara Oberbeckmann
aus Herford
2020

Mit Genehmigung der Medizinischen Fakultät der Universität München

Betreuer: PD Dr. Philipp Korber

Zweitgutachter: Prof. Dr. rer. nat. Aloys Schepers

Dekan: Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 24.7.2020

Für meine Familie, die immer an mich geglaubt hat.

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Promotionsbüro
Medizinische Fakultät

MMRS MUNICH MEDICAL RESEARCH SCHOOL

Eidesstattliche Versicherung

Anna Elisa Barbara Oberbeckmann

Name, Vorname

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Titel

Absolute nucleosome occupancy and reconstitution of nucleosome positioning mechanisms for Saccharomyces cerevisiae

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 10.11.2020

Ort, Datum

Elisa Oberbeckmann

Unterschrift Doktorandin bzw. Doktorand

TABLE OF CONTENTS

PREFACE

This thesis includes two published articles (Chapter 1 and 2), two paper manuscripts (Chapter 3 and 4). I recommend reading them in order of appearance with emphasis on Chapter 1 and 4, which contain my main work.

The first publication deals with the quite simple, although laborious aim of measuring absolute nucleosome occupancy, which to our knowledge resulted in the first high-resolution map of absolute nucleosome occupancy for an eukaryote, in this case *Saccharomyces cerevisiae*. This would have been impossible without the close collaboration with Prof. Dr. Ulrich Gerland (TU Munich) and his graduate student Michael Wolff, who performed the data analysis.

The second publication as well as the paper manuscripts deal with the long-standing, but complex subject of nucleosome positioning and array formation. In close collaboration with the group of Prof. Dr. Karl-Peter Hopfner (LMU Munich), we ventured to dissect the mechanisms of spacing activity and DNA sequence read-out by the INO80 chromatin remodeler using biochemical and structural approaches (Chapters 2-4). In Chapter 4, INO80 was compared with three additional remodeling enzymes provided by Prof. Dr. Craig Peterson's (University of Massachusetts Medical School) and Prof. Dr. Patrick Cramer's (Max-Planck-Institute for Biophysical Chemistry) groups. The studies of Chapters 3 and 4 are in revision at the Journal Nature Communications. Additionally, Chapter 4 was uploaded to BioRxiv, a pre-print server.

Detailed author contributions are listed at the beginning of each chapter.

Published work included in this thesis:

Oberbeckmann E.*, Wolff M.*, Krietenstein N., Heron M., Ellins J.L., Schmid A., Krebs S., Blum H., Gerland U., Korber P. (2019). *Absolute nucleosome occupancy map for the Saccharomyces cerevisiae genome.* Genome Res. 29, 1996–2009, DOI: https://doi.org/10.1101/gr.253419.119

Knoll K.R.*, Eustermann S.*, Niebauer V., **Oberbeckmann E.**, Stoehr G., Schall K., Tosi A., Schwarz M., Buchfellner A., Korber P., Hopfner K.P. (2018). *The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling.* Nat. Struct. Mol. Biol. 25, 823–832, https://doi.org/10.1101/gr.253419.119

Manuscripts included in this thesis:

Oberbeckmann E.*, Niebauer V.*, Watanabe S., Farnung L., Mold M., Schmid A., Cramer P., Peterson C.L., Eustermann S., Hopfner K.P., Korber P. *Ruler elements in chromatin remodelers set nucleosome array spacing and phasing.* https://doi.org/10.1101/2020.02.28.969618

Oberbeckmann E.*, Krietenstein N.*, Niebauer V., Wang Y., Schall K., Mold M., Straub T., Rohs R., Hopfner K. P., Korber P., Eustermann S. *Genome information processing by the INO80 chromatin remodeler positions nucleosomes. https://doi.org/10.1101/2020.11.03.366690*

* These authors contributed equally to the work.

SUMMARY

Nucleosomes, the basic unit of chromatin, package the genome in a repetitive and non-random way. Genome-wide nucleosome maps revealed that nucleosomes form a stereotypical pattern at actively transcribed genes. This pattern is characterized by a nucleosome depleted region (NDR) upstream of the transcription start site followed by an array of regularly spaced nucleosomes. This stereotypical NDR-array pattern is pivotal for proper transcription initiation and therefore a major regulatory element for gene expression. Additionally, nucleosome positioning plays an important role in DNA replication and DNA repair. The NDR-array pattern is to some extent encoded in the DNA sequence, which is mainly read out by a combination of non-histone DNA binding proteins (general regulatory factors, GRFs) and ATP dependent chromatin remodeling enzymes (remodelers). Deletion experiments in *Saccharomyces cerevisiae* revealed that remodelers usually have redundant functions, whereas GRFs are essential for viability, which both complicates the detailed mechanistical dissection of these proteins *in vivo*.

Therefore, the Korber group established a genome-wide remodeling assay with *in vitro*-assembled chromatin, which reconstitutes the individual contribution of each remodeling enzyme/GRF to the stereotypical NDR-array pattern. This approach revealed that some remodelers, like INO80, position *in vivo*-like nucleosomes on their own, whereas other remodelers, like ISW1a and ISW2, need an alignment point provided by GRFs. However, it remained unclear how remodelers generate nucleosome regularity in arrays and how arrays are aligned at GRFs. In particular, it was unclear to which extent remodelers generate the array-defining distances between nucleosomes, and between nucleosomes and GRFs by themselves, or if rather the nucleosome density and the underlying DNA sequence dominate these distances. Here, we showed that not just ISWI-type remodelers, but also INO80 as well as Chd1 align nucleosomes at GRFs and that all remodelers with spacing activity contain a ruler element as they generate remodeler-specific regular spacing in arrays and array alignment (phasing). This ruler most likely resides in the DNA-binding domain/subunit of each remodeler and can in some cases respond to nucleosome density. The extent of the nucleosomal arrays depends on the nucleosome density and mildly on the underlying DNA-sequence. Based on structural information of the INO80 remodeling complex, we generated INO80 mutants, which generated altered spacing and phasing

distances in our reconstitution assay. This tuned for the first-time array generation by a remodeler and revealed the location of the ruler element in INO80.

Not only the information where a nucleosome is positioned, but also how often this position is occupied, is fundamental for all nucleosome-related processes. However, all available genome-wide nucleosome mapping techniques are not able to provide nucleosome occupancy in absolute terms but rather measure nucleosome densities relative to the maximal nucleosome peak height in a single sample. To overcome this limitation, we established two orthogonal approaches to map absolute nucleosome occupancy. The first genome-wide high-resolution occupancy map of the *Saccharomyces cerevisiae* genome reveals that nucleosomal arrays exhibit uniformly high nucleosome occupancy. This contrasts other nucleosome maps, which often suggested drastic changes in nucleosome occupancy within single genes. Furthermore, we did not find any correlation between high transcription rates and low nucleosome occupancy as indicated by other studies, but we revealed a correlation between low nucleosome occupancy and high RSC occupancy, a nucleosome-evicting remodeling enzyme.

ZUSAMMENFASSUNG

Nukleosomen sind die grundlegenden Strukturelemente in Chromatin und verpacken das Genom auf eine repetitive und nicht zufällige Art und Weise. Genomweite Nukleosom-Karten zeigten, dass Nukleosomen eine stereotypische Verteilung an aktiv transkribierten Genen aufweisen. Dieses Muster ist charakterisiert durch eine nukleosomendepletierte Region (NDR) vor der Transkriptionsstartstelle gefolgt von einer Abfolge äquidistanter Nukleosomen. Dieses stereotypische NDR-Reihenmuster ist zentral für richtige Transkriptionsinitation und deshalb ein wichtiges regulatorisches Element für Genexpression. Zusätzlich spielt Nukleosomenpositionierung eine wichtige Rolle in DNA-Replikation und DNA-Reparatur. Das NDR-Reihen-muster ist teilweise in der DNA-Sequenz kodiert, welche v.a. von einer Kombination aus nicht-Histon DNA-Bindeproteinen (generelle Regulationsfaktoren, GRFs) und ATP-abhängigen Chromatinumbauenzymen (‚*Remodeler*') gelesen wird. Deletions-experimente in *Saccharomyces cerevisiae* zeigten, dass *Remodeler* normalerweise redundant arbeiten, wohingegen GRFs essenziell für die Überlebensfähigkeit der Zelle sind. Beides verkompliziert die detaillierte mechanistische Analyse dieser Proteine in vivo.

Deshalb etablierte das Korber-Labor einen genomweiten Remodeler-Assay mit In vitro-Chromatin. Dieser Assay rekonstituiert die individuellen Beiträge von jedem *Remodeler* oder GRF zu dem stereotypischen NDR-Reihenmuster. Dieser Ansatz zeigte, dass einige *Remodeler*, wie ISW1a und ISW2 aus Hefe, einen Bezugspunkt brauchen in Form von GRFs. Trotzdem blieb es unklar, wie genau *Remodeler* Regularität in Nukleosomenabfolgen erzeugen und wie diese Nukleosomenabfolgen an den GRFs ausgerichtet werden. Speziell war unklar, bis zu welchem Grad *Remodeler* die reihendefinierenden Abstände zwischen Nukleosomen und zwischen GRFs und Nukleosomen selbst einstellen oder ob eher die Nukleosomendichte und die zugrundeliegende DNA-Sequenz diese Abstände dominieren. Hier zeigen wir, dass nicht nur *Remodeler* vom ISWI-Typ, sondern auch INO80 und Chd1 aus Hefe Nukleosomen an GRFs ausrichten können und dass alle *Remodeler*, die reguläre Nukleosomenabstände erzeugen, ein strukturelles Element ähnlich einem Lineal besitzen, da sie *Remodeler*-typische reguläre Abstände innerhalb Nukleosomreihen oder zwischen Nukleosomreihen und Ausrichtungspunkt erzeugen. Dieses Lineal-Element liegt wahrscheinlich in der DNA-binde-Domäne/Untereinheit des einzelnen *Remodelers* und in manchen Fällen reagiert dieses Lineal auf die Nuklesomendichte.

Das Ausmaß der Nukleosomreihung hängt hauptsächlich von der Nukleosomendichte ab und teilweise von der DNA-Sequenz. Basierend auf strukturellen Daten des INO80-Komplexes konnten wir INO80-Mutanten erzeugen, welche veränderte Abstände in unserem Rekonstitutionssystem einstellten. So gelang zum ersten Mal die Manipulation der Bildung von Nukleosomreihen durch einen *Remodeler*. Zudem identifizierte es die Lage des Lineal-Elementes in INO80.

Nicht nur die Information, wo ein Nukleosom positioniert ist, sondern auch wie oft es diese Position besetzt, ist fundamental für alle nukleosomen-abhängigen Prozesse. Nichtsdestotrotz sind alle verfügbaren genomweiten Methoden zur Kartierung von Nukleosomen nicht in der Lage, den Nukleosomenbesetzungsgrad vollständig zu messen. Stattdessen messen diese Methoden eher eine Nukleosomendichte, die relativ zur maximalen Nukleosomenbesetzung ein jeder Probe berechnet wird. Um dieses Problem zu überwinden, haben wir zwei orthogonale Methoden entwickelt, die absolute Nukleosomenbesetzung messen. Die daraus resultierende erste genomweite, hochauflösende Nukleosomen-Besetzungskarte für das Saccharomyces cerevisiae Genom zeigt Nukleosomreihen mit gleichmäßig hoher Nukleosomenbesetzung. Das steht im Gegensatz zu anderen Nukleosomkarten, die oft einen zweifachen Unterschied zwischen Nukleosomenbesetzung in demselben Gen suggerieren. Des Weiteren konnten wir keine Korrelation zwischen hohen Transkriptionsraten und niedriger Nukleosomenbesetzung feststellen, obwohl andere Studien darauf hinweisen. Jedoch konnten wir eine Korrelation zwischen niedriger Nukleosomenbesetzung und dem vermehrten Vorkommen des *Remodelers* RSC ableiten, welcher Nukleosomen entfernt.

INTRODUCTION

Eukaryotic DNA is packaged into a structure called chromatin, which was first identified and named by the cytogeneticist W. Flemming in the late 19th century (Flemming, 1882) (Figure 1A). He discovered a substance in the cell nucleus that was readily stained and named it after the ancient greek word *chroma,* which means color.

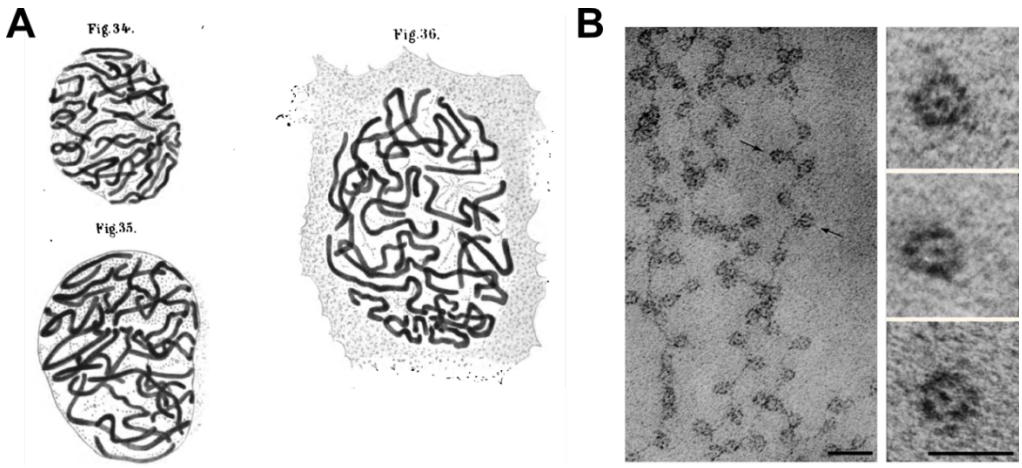


Figure 1. First visualizations of chromatin. **(A)** Drawings of *Salamandra maculosa* epithelium cells stained with chromic acid. Black spaghetti correspond to chromosomes in a condensed form of chromatin. [From W. Flemings '*Zellsubstanz, Kern und Zelltheilung*', 1882.] **(B)** Chromatin in form of 'Beads-on-a-string' (left) or mono-nucleosomes (right) visualized by electron microscopy. Size marker: 30 nm (left), 10 nm (right). [From Olins and Olins, 2003.]

Chromatin comprises mainly nucleic acids and small proteins called histones as described already by F. Miescher, A. Kossel and others in the late 19th century (Olins and Olins, 2003). It took nearly anorther century to discover the basic repeating unit of chromatin: the nucleosome. A. and D. Olins, and C. Woodcock could visualize by transmission electron microscopy that chromatin consists of repetitive particles (nucleosomes), which are ordered like beads-on-a-string, (Olins and Olins, 1974; Woodcock et al., 1976) (Figure 1B). Based on this, nuclease digestion data and crosslinking data, R. Kornberg proposed that the nucleosome consists of ~200 bp of DNA in complex with four histone dimers (Hewish and Burgoyne, 1973; Kornberg, 1974; Kornberg and Thomas, 1974). In 1975, Oudet and colleagues termed these particles nucleosomes (Oudet et al., 1975).

Chromatin organization

The nucleosome

The nucleosome is defined as the nucleosome core particle (NCP) plus ~10-80 bp of linker DNA depending on cell type and species (Holde, 1989). The NCP comprises 147 base pairs (bp) of DNA wrapped in 1.65 left-handed super-helical turns around a histone octamer, which contains two copies of each core histone H2A, H2B, H3 and

H4 (Figure 2A) (Davey et al., 2002; Luger et al., 1997). The pseudo-twofold axis of the NCP defines the center of the 147 bp and is called nucleosome dyad. Starting bidirectionally from the dyad, super-helical locations (SHL) -7 to 7 are defined in 10 bp steps. The relatively small core histones (11-15 kDa) are positively charged to facilitate binding and bending of the negatively charged DNA (Clark and Kimura, 1990). They assemble into four heterodimers (two each of H2A/H2B, H3/H4) and all together form a spool-like structure, which tightly binds the core DNA. The disc-like NCP is about 5.5 nm high and 11 nm in diameter. Each histone contains a central histone-fold motif comprised of three α-helices (Figure 2B), while the N-terminus, and for H2A also the C-terminus, is mainly unstructured ("histone tails") and often target of epigenetic modifications (Peterson and Laniel, 2004; Zhao and Shilatifard, 2019)

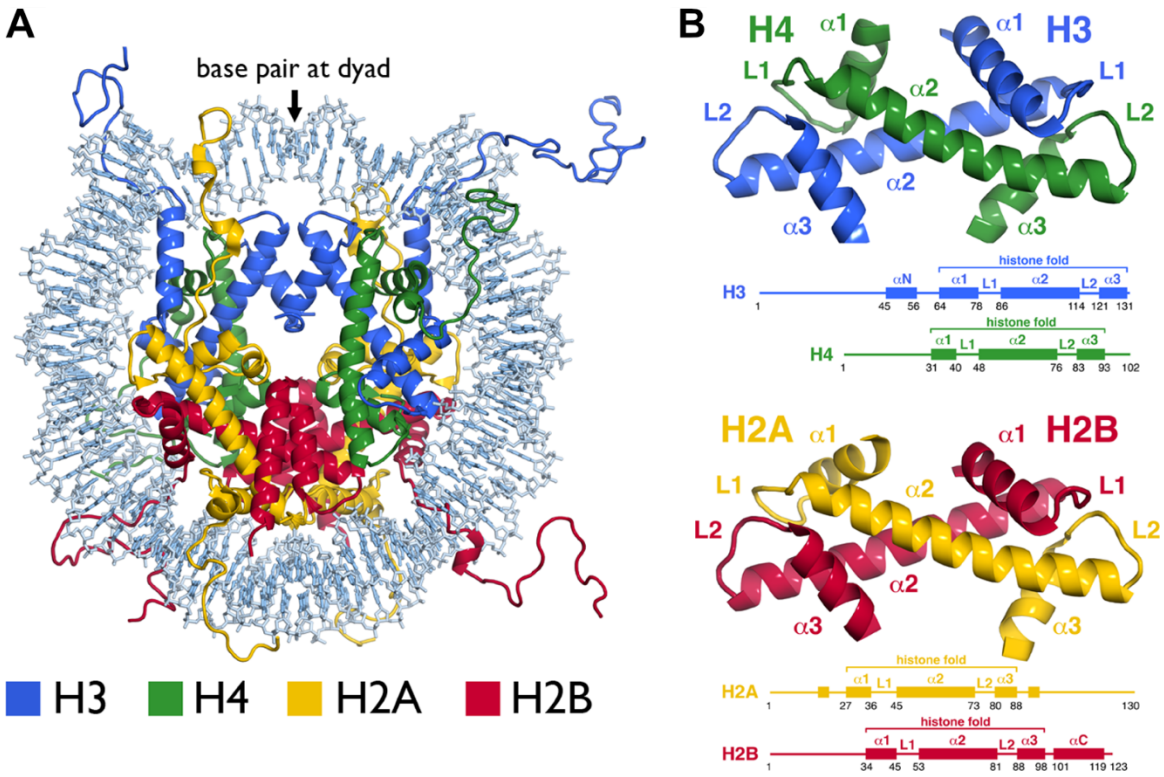


Figure 2. The nucleosome core particle. **(A)** High resolution structure of the nucleosome core particle consisting of recombinant *Xenopus laevis* histones and human alpha-satellite DNA (Davey et al., 2002). **(B)** Histone-fold structure of the four core histones as heterodimers H3/H4 (top) and H2A/H2B (bottom). Below: details of the secondary elements of each histone are indicated (PDB ID 1KX5). [From McGinty and Tan, 2015.]

Linker histone and histone variants

A fifth histone type, called H1/H5, binds about 20 bp of linker DNA near the nucleosome dyad. Together, nucleosome and linker histone are called the chromatosome (Simpson, 1978). The linker histone stabilizes the nucleosome (Clark and Kimura, 1990) and, in higher eukaryotes, it is often associated with transcriptionally inactive chromatin, also called heterochromatin. In flies, for example, 30% of the genome

constitutes heterochromatin, which contains very repetitive and gene-poor DNA sequences (Smith et al., 2007). However, in *Saccharomyces cerevisiae* 85% of the genome is expressed (David et al., 2006) and only minor parts, like the mating type locus and telomers, comprise silenced heterochromatin (Bi, 2014; Grunstein, 1997). Additionally, linker histones are less conserved than the highly-conserved core histones, and the *S. cerevisiae* H1 homolog Hho1 is neither essential nor very abundant (0.05 molecules Hho1/nucleosome) (Lawrence et al., 2017; Patterton et al., 1998). Besides the five canonical histones, *S. cerevisiae* contains two histone variants: the non-essential H2A.Z (Htz1), which is important for proper gene expression and genome stability (Marques et al., 2010), and the essential Cse4, which replaces H3 in centromers (Meluh et al., 1998). Higher eukaryotes have a larger variety of histone variants that extend the epigenetic variability of chromatin (Talbert and Henikoff, 2017; Zink and Hake, 2016).

Higher-order chromatin structures

Heterochromatin and linker histones are often associated with condensed chromatin. In contrast to the well-established 10-nm fiber (Figure 3), the form and extent of higher-order structures, like the 30-nm fiber, are still debated (Fussner et al., 2011; Maeshima et al., 2019, 2016, 2014, 2010). Initially, a 30-nm chromatin fiber was identified by transmission electron microscopy (Finch and Klug, 1976). A first model was proposed which was called ‘solenoid’ or ‘one-start helix’ where nucleosomes are consecutively located next to each other (Figure 3). Later, the ‘zigzag’ or ‘two-start helix’ model was proposed, where every other nucleosome is stacked in a zigzag manner onto each other (Figure 3) (Woodcock et al., 1984). The ‘zigzag’ model was supported *in vitro* by cryoEM structures (Garcia-Saez et al., 2018; Song et al., 2014).

However, more and more evidence accumulated that the 30-nm fiber is only formed *in vitro* under low-salt conditions or *in vivo* in specific cell types (e.g. chicken erythrocytes or starfish sperm) (Hansen, 2012; Joti et al., 2012; Maeshima et al., 2014). Interestingly, these cell types often contain high H1 concentration, for example: ~1.4 H5 molecules/nucleosome in chicken erythrocytes or ~1.7 H1 molecules/nucleosome in starfish sperm (Kowalski and Pałyga, 2011), in contrast to somatic cells with ~0.65 H1 molecules/nucleosome (Woodcock et al., 2006). Nonetheless, many other studies could not find evidence for a 30-nm fiber *in vivo*, for example by electron microscopy (McDowall et al., 1986), fluorescent microscopy (Ricci et al., 2015), small-angle X-ray scattering (SAXS) (Nishino et al., 2012) or chromosome-conformation-capture (3C)

(Dekker, 2008). Instead, they described a more dynamic, liquid-like chromatin structure consisting of irregularly folded 10-nm fibers, which can be easily assessed and regulated, and called it ‘polymer melt’ or ‘nucleosome clutches’. Still, specific regions or short stretches in the genome might form 30-nm fibers.

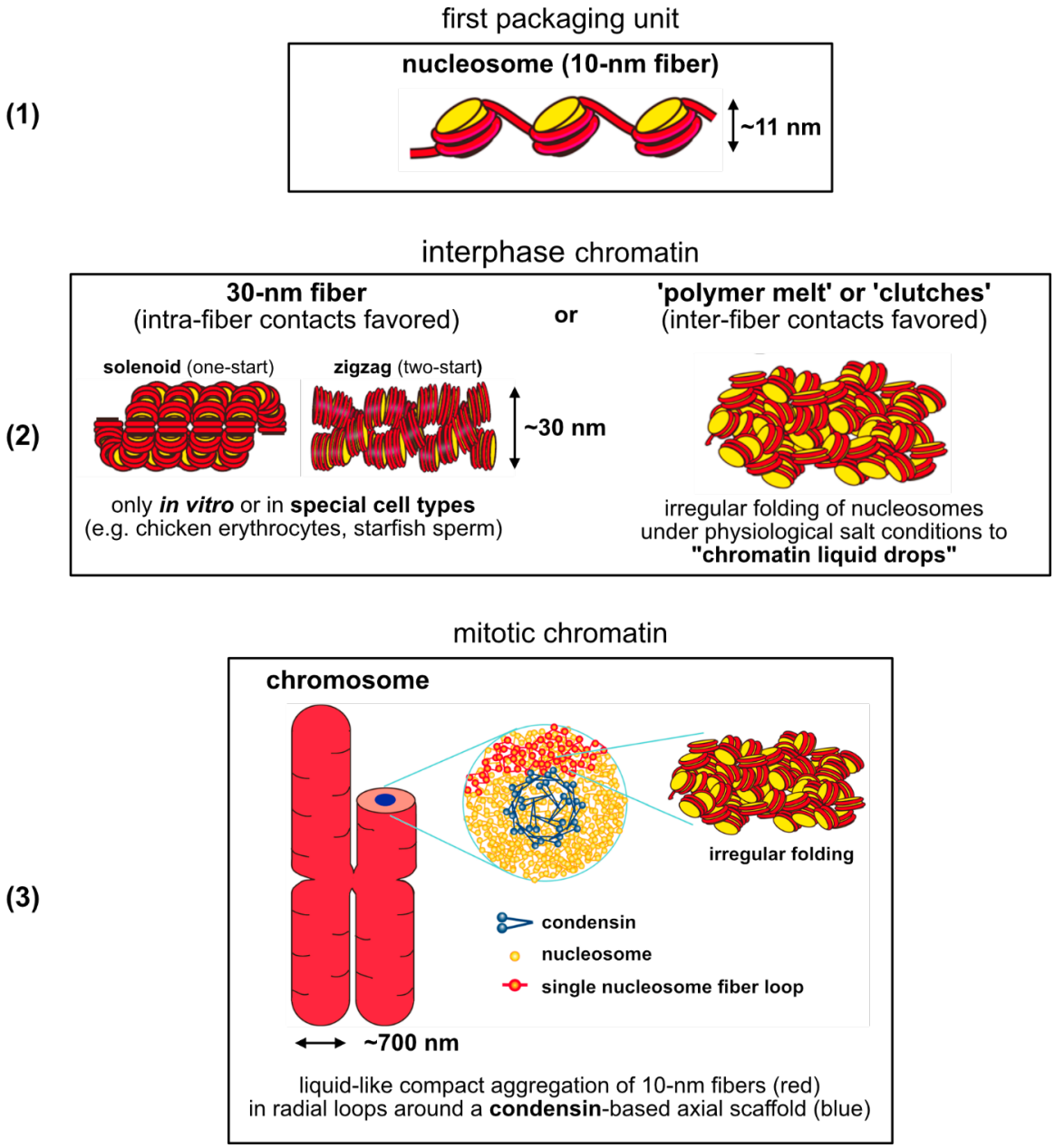


Figure 3. Levels of chromatin compaction. **Level (1)** 10-nm fiber consisting of nucleosomes. **Level (2)** Structured 30-nm fiber versus ‘polymer melt’ or ‘clutches’ of nucleosomes. The 30-nm fiber is formed when contacts within one fiber are favored, what is the case in low salt conditions e.g. *in vitro*. At physiological salt conditions contacts between several different nucleosome fibers are usually preferred, what leads to nucleosome clutches or a polymer melt. **Level (3)** Higher-order structures in mitotic chromatin. [Modified from Maeshima et al., 2014, 2016.]

During mitosis, interphase chromatin is compacted further such that individual chromosomes are detectable by light microscopy (Figure 3). Again, several mechanisms were proposed how chromatin folds into mitotic chromosomes. The hierarchical folding model proposes that the 30-nm fiber is compacted to a 100-nm

fiber, 200-nm fiber and so on by hierarchical folding (Sedat and Manuelidis, 1978). However, more recent studies showed that chromatids have a condensin-based axial scaffold which is radially surrounded by irregularly folded 10-nm fibers loops formed by a loop extrusion mechanism (Figure 3) (Mirny et al., 2019).

Stereotypical nucleosome organization and its implications for transcription

Nucleosomes as genomic regulators

The nucleosome is intrinsically obstructive and therefore a major regulatory element for fundamental processes like transcription, replication or DNA repair (Bell et al., 2011; Lai and Pugh, 2017; Radman-Livaja and Rando, 2010). For example, transcription factors forming the preinitiation complex (PIC) during transcription need to bind promotor DNA, which must be nucleosome-free. Therefore, nucleosomes can function as regulatory elements for promoter activation as shown *in vivo* (Almer et al., 1986; Han and Grunstein, 1988; Fascher et al., 1990), and can be inhibitory for transcription itself as shown *in vitro* with chromatinized plasmids (Knezetic and Luse, 1986; Lorch et al., 1987; Workman and Roeder, 1987). Thus, mapping of *in vivo* nucleosome positions to understand the regulatory power of nucleosomes is of major interest, and in the last decades it became feasible for whole genomes.

Mapping nucleosomes

Limited endonuclease digests of chromatin showed that DNA within the NCP is protected against nucleases leading to regularly spaced DNA bands (“ladders”) in gel electrophoresis (Figure 4A) (Clark and Felsenfeld, 1971; Hewish and Burgoyne, 1973; Noll, 1974). This approach was refined by using micrococcal nuclease, an endo-exonuclease from *Staphylococcus aureus* (Telford and Stewart, 1989). The regularity of the ladder pattern can be quantified as the average nucleosome repeat length (NRL), which is 165 bp in *S. cerevisiae* (147 bp NCP + 18 bp linker) (Holde, 1989). The NRL varies between species (~154 bp in *S. pombe* (Lantermann et al., 2010), ~200 bp in flies), between cell types (~173 bp in rat cortical neurons,~199 bp in rat liver cells), and between transcriptional states (160 bp versus 170 bp for genes in *S. cerevisiae* with high versus low transcription rate, respectively (Chereji et al., 2018)).

Not just the average NRL, but individual nucleosome positions relative to the genome sequence can be determined by high-throughput sequencing of the protected mononucleosomal DNA fragments after MNase digests (Albert et al., 2007; Shivaswamy et al., 2008; Tirosh et al., 2010). This so-called MNase-seq method can

gain specificity by immunoprecipitation of specific histones or histone variants, e.g. H2A.Z (Albert et al., 2007; Mavrich et al., 2008a, 2008b). This adaptation was named MNase-ChIP-seq (Micrococcal Nuclease-Chromatin Immunoprecipitation-sequencing). An important and precise alternative to MNase-seq is chemical cleavage mapping (Brogaard et al., 2012), and especially its recent improved version (Chereji et al., 2018). For precise cleavage, a cysteine gets introduced into one of the histones, which leads to DNA cleavage close to this cysteine upon addition of iodoacetamide-coupled phenanthroline and hydrogen peroxide.

Stereotypical NDR-array organization at promoters

Genome-wide nucleosome maps revealed that nucleosomes are not randomly distributed but form a stereotypical pattern around transcription start sites (TSS) (Figure 4B). The region upstream of the TSS is nucleosome depleted and called nucleosome-depleted or -free region (NDR or NFR, respectively). It is flanked by two well positioned nucleosomes called -1 and +1 nucleosome followed by regularly spaced nucleosomes that form phased nucleosomal arrays. Arrays are phased if their nucleosomes have the same regular spacing and their start points are aligned at common a reference point, e.g., the TSS, over a set of genes or a population of cells.

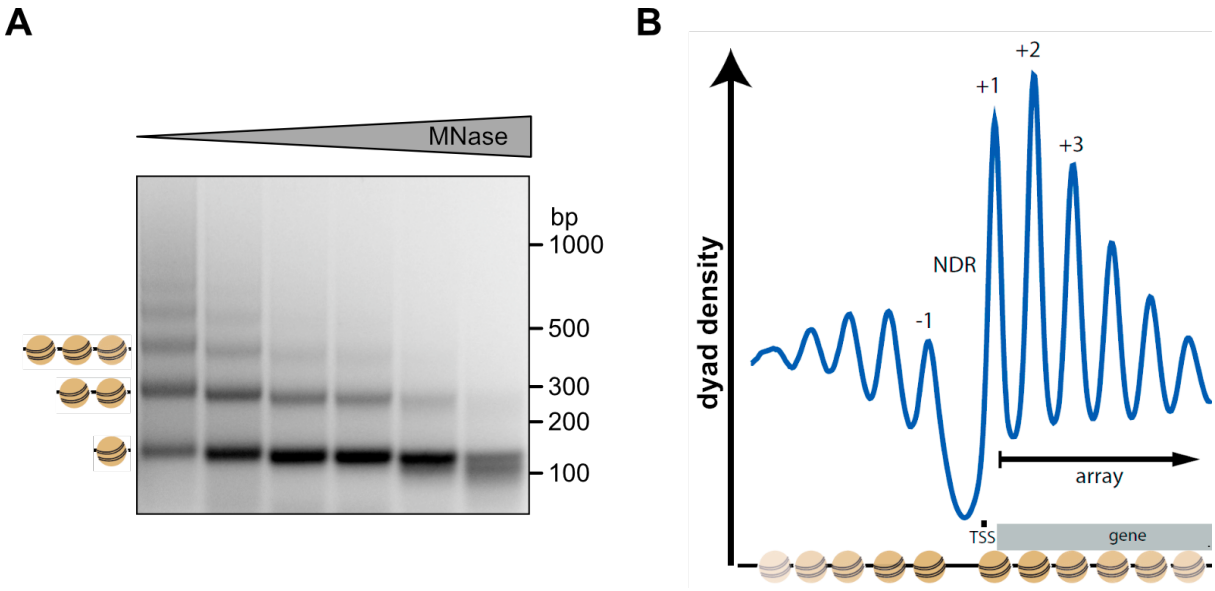


Figure 4. Regularity in chromatin visualized by MNase-digestion. **(A)** Micrococcal nuclease (MNase) digestion of *S. cerevisiae* chromatin with MNase concentration increasing from left to right. Purified DNA after MNase digests was subjected to agarose gel electrophoresis and stained with ethidium bromide. Kindly provided by A. Schmid (Korber group). **(B)** Stereotypical NDR-array pattern derived from *S. cerevisiae* MNase-seq data. Extended nucleosomal dyad densities aligned at transcription start sites (TSS) were averaged over all annotated genes. [From Lieleg et al., 2015.]

In yeast, the TSS is located 12-13 bp within the first nucleosome (Figure 4B, Bean et al., 2016; Lantermann et al., 2010). It is still unclear, if this is cause or consequence of

transcription initiation *in vivo*, but *in vitro* studies showed that +1 nucleosome positioning is mainly independent of transcription in *S. cerevisiae* (Krietenstein et al., 2016; Zhang et al., 2011). This points towards an establishment of the +1 nucleosome position first, which consequently would dictate the transcription initiation site (Jiang and Pugh, 2009). In higher eukaryotes, the TSS is located upstream of the +1 nucleosome (Bai and Morozov, 2010; Mavrich et al., 2008b; Schones et al., 2008).

Transcription activation requires promoter NDRs

Although the stereotypical NDR-array pattern at TSSs seems to be omnipresent in yeast, it is mainly established at housekeeping (growth genes) or constitutively active genes but not at silent or inducible genes. This becomes more apparent in higher eukaryotes, where more genes are silent (Mavrich et al., 2008b; Schones et al., 2008; Tirosh and Barkai, 2008). Therefore, genes can be subdivided into canonical (stereotypical) and non-canonical (non-stereotypical) genes, which either show the stereotypical NDR-array pattern at their TSSs or not, respectively. Non-canonical promotor architecture differs, e.g., the NDR is narrowed, shifted upstream, or filled with nucleosomes. Prominent yeast examples are *PHO* and *GAL* promotors. Upon induction by phosphate starvation or galactose availability, 3-4 nucleosomes get disrupted, what leads to the availability of regulatory elements like the TATA box and results in transcription activation (Lohr, 1997).

Determinants of nucleosome positioning in *S. cerevisiae*

To understand how eukaryotic cells regulate DNA accessibility by nucleosomes, researchers started exploring which factors determine nucleosome positioning.

Cis-factor: DNA sequence-intrinsic nucleosome positioning

First, the DNA sequence was investigated as a direct intrinsic positioning factor (cis-factor) (Kaplan et al., 2009; Segal et al., 2006). *In vitro* reconstitution by salt gradient dialysis (SGD) showed that long poly(dA:dT) stretches can repel nucleosomes which seemed particularly relevant for *S. cerevisiae,* where promoter NDRs are rich in poly(dA:dT) stretches (Anderson and Widom, 2001; Iyer and Struhl, 1995; Zhang et al., 2011). Furthermore, eukaryotic NCP sequences revealed a dinucleotide periodicity. Every 10 bp either an AA/TT/TA dinucleotide or a GC dinucleotide with 5 bp offset to the A/T dinucleotides was found (Albert et al., 2007; Brogaard et al., 2012; Chereji et al., 2018; Segal et al., 2006). This dinucleotide periodicity favors appropriate widening and compression of the DNA helix grooves resulting in lower DNA bending

energies. Dinucleotide periodicity can be observed in many eukaryotes to different extents, but not in prokaryotes (Bettecken and Trifonov, 2009; Zhang et al., 2009).

Trans-factors: GRFs and chromatin remodelers

DNA-encoded effects could not explain all nucleosome positions found *in vivo,* and soon it became clear that trans-factors modulate nucleosome positioning (Zhang et al., 2009, 2011). Among them: non-histone DNA binding proteins, also called general regulatory factors (GRFs) in yeast, and a family of SNF2-type helicases, called ATP dependent chromatin remodeling enzymes (remodelers).

GRFs are sequence-specific DNA binders, often essential for cell viability and have diverse functions, for example, in transcription termination, chromatin silencing or DNA repair (Chen et al., 2011; Colin et al., 2014; Fourel et al., 2002; Reed et al., 1999). They have in common to antagonize nucleosome deposition at their specific DNA binding motifs, which are located mainly in promotor regions or other NDRs (Badis et al., 2008). This was shown on genome scale by degron-induced depletion of the yeast GRFs Abf1 or Reb1, which resulted in nucleosome-occupied GRF-binding sites genome-wide (Hartley and Madhani, 2009). Bai and colleagues classified yeast Abf1, Cbf1, Mcm1, Rap1, Reb1, and Orc1 as strong nucleosome depletion factors *in vivo* by cloning their respective DNA-binding motifs into nucleosome-occupied positions, which turned these into NDRs (Yan et al., 2018). Furthermore, GRFs can function as barrier for aligning nucleosomal arrays as shown *in vitro* (Krietenstein et al., 2016; Li et al., 2015). This alignment is mediated by ATP dependent chromatin remodeling enzymes, short “remodelers”, since GRFs have no ATPase activity and cannot translocate nucleosomes by themselves as shown *in vitro* (Krietenstein et al., 2016).

Remodelers mobilize or alter nucleosomes by breaking DNA-histone contacts and translocating DNA upon ATP hydrolysis (Zhou et al., 2016). They can be monomeric but are usually part of large multisubunit complexes, and belong to the family of SNF2-type helicases (Flaus et al., 2006), which share two conserved RecA-like ATPase lobes as main catalytic subunit (Table 1). The two lobes are asymmetrical. While lobe 1 contains a Walker A and B motif, which binds and hydrolyzes ATP, respectively, lobe 2 contains an arginine finger motif instead. Closure of both lobes is required for proper coupling of ATP hydrolysis to DNA translocation. Depending on sequence similarity of the catalytic subunit, SNF2-type proteins can be classified into four subfamilies: SWI/SNF, CHD, ISWI, and INO80. This homology sometimes results in similar

functional properties: disassembly of nucleosome (eviction), movement of nucleosomes (sliding), setting regular distances between nucleosomes (spacing) or exchanging histones for histone variants (histone exchange) (Table 1).

Table 1. Overview of the most prominent remodeling enzymes in *S. cerevisiae*. Catalytic subunit refers to the main ATPase exhibiting DNA translocase activity.

subfamily	remodeling complexes	remodeler functions	catalytic subunit	number of subunits
SWI/SNF	RSC	sliding, eviction	Sth1	15-16
	Swi/Snf	sliding, eviction	Swi2/Snf2	8-14
ISWI	ISW1a	sliding, spacing	Isw1	2
	ISW1b	sliding	Isw1	3
	ISW2	sliding, spacing	Isw2	2
CHD	Chd1	sliding, spacing	Chd1	1
INO80	INO80	sliding, spacing, Htz1 removal	Ino80	15
	SWR1	Htz1 incorporation	Swr1	14

Remodeler subfamily members in S. cerevisiae

RSC and Swi/Snf complexes can slide and disassemble nucleosomes (Clapier and Cairns, 2009). Both complexes have homologous ATPases, and several homologous or even identical subunits. However, RSC is essential and ten times more abundant than SWI/SNF (Cairns et al., 1996; Ghaemmaghami et al., 2003). Furthermore, RSC reads poly(dA:dT) tracts, which leads to nucleosome eviction at poly(dA:dT)-rich NDRs in *S. cerevisiae* (Krietenstein et al., 2016; Lorch et al., 2014). Therefore, RSC is one of the main drivers of NDR formation at promoters in budding yeast (Hartley and Madhani, 2009; Krietenstein et al., 2016; Kubik et al., 2019; Parnell et al., 2008).

The ISWI family contains three remodeling complexes with two different catalytic subunits: Isw1 and Isw2 (Table 1). Isw2 associates with Itc1, whereas Isw1 associates either with Ioc3 (ISW1a) or Ioc2 and Ioc4 (ISW1b) (Vary et al., 2003). Although ISW1a and ISW1b share the same catalytic subunit, their remodeling activities are different: ISW1a can slide and space nucleosomes, while ISW1b can only slide nucleosomes (Stockdale et al., 2006). ISW2 also exhibits spacing activity, and further it is involved in +1 nucleosome positioning (Kubik et al., 2019), but less active in gene bodies (Gkikopoulos et al., 2011; Ocampo et al., 2016) as *in vivo* studies suggested.

Chd1, the only member of the CHD family in budding yeast, exhibits spacing activity (Stockdale et al., 2006), and deletion studies suggested that Chd1 generates a tighter

NRL (~160 bp) than ISW1a (~175 bp) or ISW2 (~200 bp) (Ocampo et al., 2016). Together, Chd1 and ISW1a, are the main drivers for array formation and spacing *in vivo*, and only deletion of both remodeling enzymes disturbed array formation significantly. Chd1 probably associates with RNA polymerase II and elongation factors like PAF, Spt5 or FACT (Simic et al., 2003). Therefore, Chd1 is most likely recruited to transcribed genes, whereas ISW1a may act more globally.

The INO80 family is characterized by a long insertion between lobe1 and lobe 2 of the main ATPase. Structural studies showed that two AAA^+ ATPase subunits, Rvb1 and Rvb2, form a hexameric ring structure, which engulfs the insert between the ATPase domains like a chaperone, and further serves as a scaffold for the assembly of other subunits (Eustermann et al., 2018; Willhoft et al., 2018). The SWR1 complex exchanges an H2A-H2B dimer for an H2A.Z-H2B dimer mainly at +1 nucleosomes but lacks sliding activity (Mizuguchi et al., 2004). In contrast, the INO80 complex can slide and space nucleosomes (Udugama et al., 2011; Zhou et al., 2018), and might be responsible for the reverse exchange of H2AZ-H2B with H2A-H2B (Brahma et al., 2017; Papamichos-Chronakis et al., 2011). However, INO80's histone exchange activity is still debated (Wang et al., 2016).

Yeast *Ino80* deletion mutants are sensitive to DNA damage agents suggesting a role of INO80 at DNA double strand breaks (Shen et al., 2000). Furthermore, INO80 is involved in +1 nucleosome positioning probably by intrinsic reading of DNA sequence information encoded at some +1 positions (Krietenstein et al., 2016; Kubik et al., 2019).

Transcription and nucleosome positioning

Besides remodeling enzymes and GRFs, transcription and especially transcription elongation might have effects on nucleosome positioning. In *S. cerevisiae*, RNA polymerase II depletion leads to wider nucleosome spacing, and in case of temperature sensitive RNA polymerase II mutants to a downstream shift of the +1 nucleosome (van Bakel et al., 2013; Weiner et al., 2010). This correlates with short spacing at highly transcribed genes (Chereji et al., 2018). However, it is unclear if this is directly due to the elongating polymerase or rather to the co-recruited Chd1 remodeler, which could set a shorter spacing after/during polymerase passage.

Integrative model for NDR-array formation

In summary, the stereotypical NDR-array pattern is generated by the combined action of several factors, which influence nucleosome positioning to different degrees (Figure

5). First, the NDR is created by a combination of nucleosome repelling poly(dA:dT) elements, GRFs, and remodelers like RSC, which probably interacts with both. Second, the +1 nucleosome is positioned most likely by INO80 via direct DNA sequence read-out or by ISW2 in combination with GRFs. Finally, the array is generated by spacing enzymes like ISW1a and Chd1.

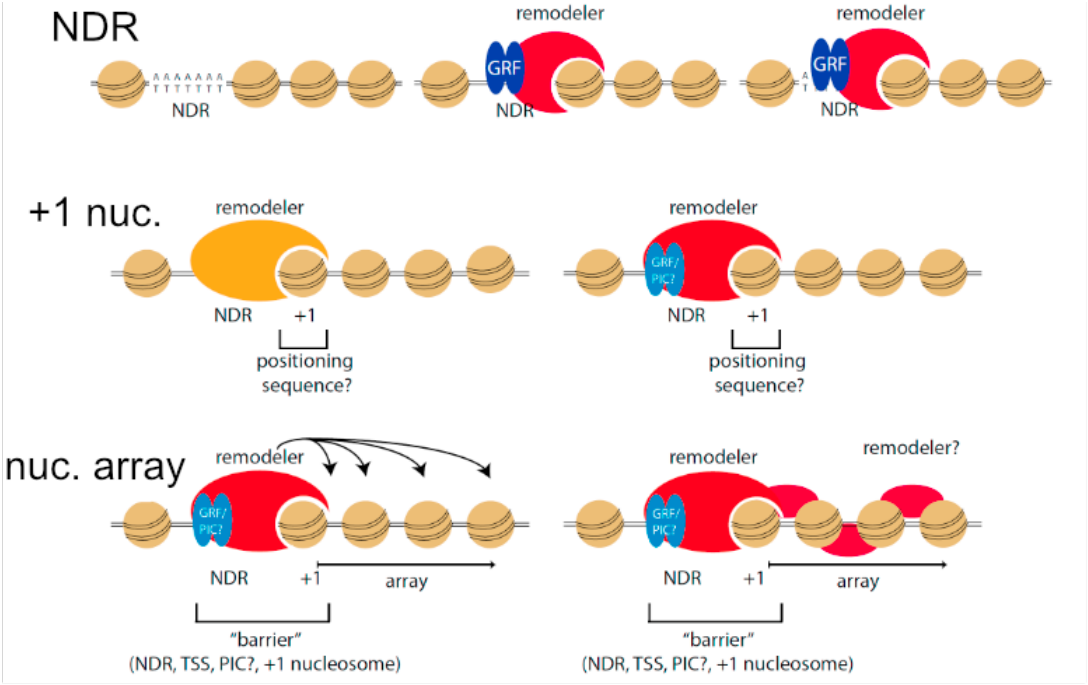


Figure 5. Integrative model summarizing different mechanisms leading to a stereotypical NDR-array organization. Formation of NDR, +1 nucleosome and nucleosomal arrays by cis- (DNA sequence) and trans-factors (GRFs, remodelers and eventually PIC). [Modified from Lieleg et al., 2015.]

Open question: mechanism of nucleosome spacing

Still unsolved is the question how exactly remodeling enzymes set regular distances between nucleosomes and how they recognize an alignment point (GRF, DNA sequence etc.) for the generation of phased arrays, and whether there is a length difference between nucleosome spacing and aligning nucleosomes against barriers. Several models were proposed to explain nucleosome packaging mechanisms, for example the statistical positioning model or the length sensor model. Statistical positioning implies that array formation is the sum of statistical interactions between nucleosomes relative to a barrier (Kornberg and Stryer, 1988), whereas the length sensor model suggests that sliding activity depends on linker length leading to linker length equalization (Yang et al., 2006; Zhou et al., 2018). In both models, linker length would reciprocally depend on nucleosome density. However, *in vitro* experiments with decreased nucleosome density in combination with whole cell extracts from *S. cerevisiae* (Zhang et al., 2011) or purified remodelers (Lieleg et al., 2015a) showed

that spacing can remain constant. Similar results were observed *in vivo* with histone-depletion strains (Gossett and Lieb, 2012; van Bakel et al., 2013). Concordant with such density-independent spacing would be the protein ruler model (Yamada et al., 2011). This model is based on structural studies on ISW1a, and suggested that the spacing activity of this remodeler may employ a DNA binding domain/subunit, which sets a fixed distance between nucleosomes. Further, this model entails that the remodeler pulls a nucleosome towards the neighboring nucleosome, because the protein ruler needs to be located between both nucleosomes to set the linker length. This would agree with the recent suggestion that some remodelers are “pullers” (INO80, ISW2) and some are “pushers” (RSC, Swi/Snf) (Kubik et al., 2019), although this terminology was coined in relation to promoter NFRs and not in relation to neighboring nucleosomes. While highly suggestive, the protein ruler model needs validation and quantification for ISW1a, and possibly other remodeling enzymes with spacing activity.

Studying stereotypical chromatin organization in vitro

Although many *in vivo* studies tried to unravel the specific contributions of remodeling enzymes by remodeler gene deletion or remodeler protein depletion (Kubik et al., 2019; Ocampo et al., 2016; van Bakel et al., 2013), it is still difficult to conclude from these studies if the remodeler contribution is direct or indirect, and how stringently individual remodelers are required and what their exact mechanistic contribution is, since remodeling enzymes have redundant functions and other nuclear processes like transcription or replication may further influence nucleosome positioning. Problems also arise with regard to dissecting GRF contributions as these are essential for viability, what restricts stringent genetic analyses. Therefore, the Korber group established a bottom-up *in vitro* approach, in which the direct mechanistic contribution of each remodeling enzyme or GRF can be tested individually on a genome-wide level (Krietenstein et al., 2012). Chromatin is assembled *in vitro* by performing salt gradient dialysis (SGD) with a plasmid library containing the *S. cerevisiae* genome and histone octamers. Incubation of SGD chromatin with whole-cell extracts from budding yeast showed that the *in vivo*-like NDR-array pattern could be reconstituted in an ATP-dependent manner (Zhang et al., 2011). By using pure proteins only, this approach could dissect that some remodelers like RSC and INO80 intrinsically read out DNA-sequence information, whereas others like ISW1a and ISW2 need an alignment point, in form of GRFs, to generate NDR-array patterns. Based on this study, new questions

arose, for example, how remodelers like INO80 intrinsically read out DNA sequence information to position nucleosomes or why Chd1 did not show spacing activity in these assays despite indications for Chd1's role in nucleosome positioning from *in vivo* studies (Gkikopoulos et al., 2011; Ocampo et al., 2016).

Absolute nucleosome occupancy

Nucleosome positions always go hand in hand with nucleosome occupancies, which describe how often a nucleosome is located at a certain genome position in a population of cells. Here, absolute nucleosome occupancy is defined as the probability of a base pair at a certain genome position to be within any NCP (Figure 6A).

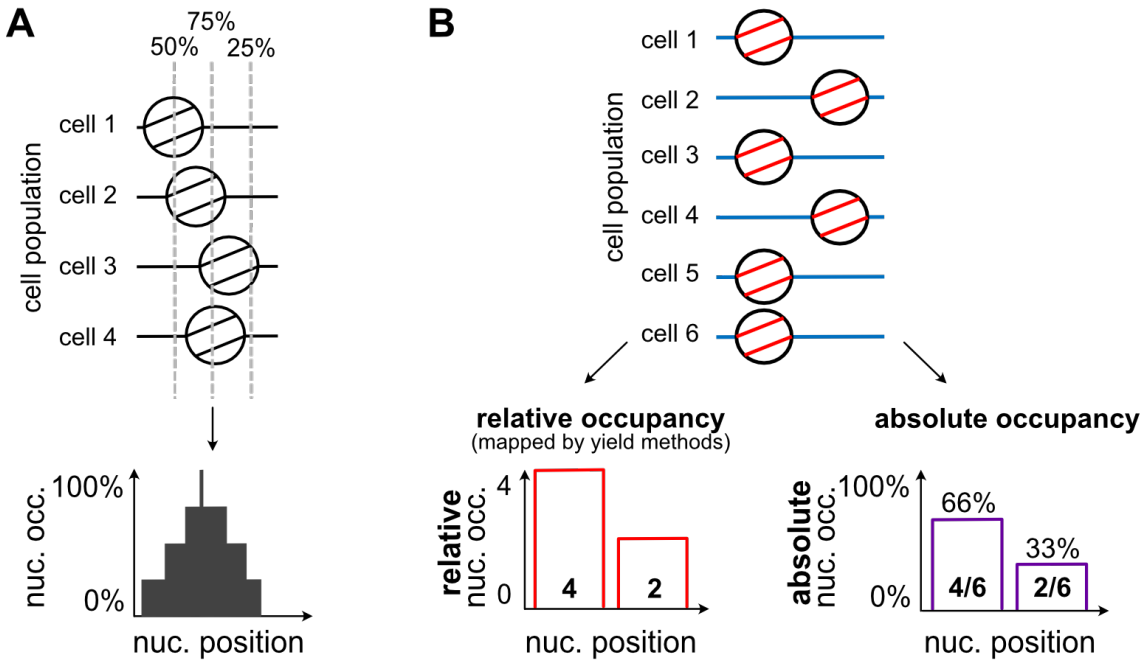


Figure 6. Definition and types of nucleosome occupancy. **(A)** Nucleosome occupancy describes the probability of a given base-pair to be within any nucleosome. **(B)** Depending on the readout, either relative or absolute nucleosome occupancy can be mapped. Relative occupancy measures only the number of fragments mapped at a certain position relative to fragments mapped to another position. In contrast, absolute nucleosome occupancy would give a complete picture of occupied and free nucleosome positions.

Current genome-wide nucleosome mapping cannot provide absolute occupancies

So far, all genome-wide nucleosome mapping techniques, if at all, only determine relative nucleosome occupancies, i.e., how many NCPs were scored at one position relative to another position in the same sample (Figure 6B). We and others called these methods yield methods, since they either map *only* nucleosome-occupied DNA (MNase-seq, chemical cleavage mapping etc.) or *only* nucleosome-free DNA (ATAC-seq (Buenrostro et al., 2013), FAIRE-seq (Giresi et al., 2007)) and therefore cannot provide the fraction of a total. In addition, MNase-based methods have several more limitations. First, MNase cleavage is 30-fold more likely at adenine and thymidine

nucleotides (Dingwall et al., 1981; Hörz and Altenburger, 1981). Second, MNase can overdigest chromatin leading to a loss of nucleosomal fragments (Figure 4A, right-most lane). Consequently, limiting digestion degrees have to be chosen, and arbitrarily chosen digestion degree result in different relative nucleosome occupancies, especially around the +1 nucleosome (Chereji et al., 2017; Weiner et al., 2010). To obtain reliable relative nucleosome occupancies, MNase-digestion degrees have to be titrated and DNA fragment yield normalized by spike-in controls. A method considering these requirements was developed only very recently (q-MNase-seq, Chereji et al., 2019) and thus, is not part of the standard MNase protocols.

Absolute nucleosome occupancy determination

Genome-wide absolute nucleosome occupancy is so far a missing quantity in chromatin biology, but necessary for correct correlations of nucleosome organization with features like transcription rates. Further, absolute nucleosome occupancy is linked to the chemical potential (concentration) of nucleosomes, which is required for modelling nucleosome distributions, e.g., in a Tonks gas model (Möbius and Gerland, 2010). To determine absolute nucleosome occupancy, it is necessary to map either nucleosome-occupied and nucleosome-free DNA simultaneously, or all nucleosomes of single cells followed by averaging over the population. The latter is up to date technically not feasible due to poor single cell genome coverage. Regarding the former option, several single-locus studies are available (Almer et al., 1986; Barbaric et al., 1992) that employ restriction enzymes (REs), which cannot cut nucleosome-occupied DNA. Therefore, the ratio of cleaved versus uncleaved DNA fragments, as determined for example via Southern blotting, corresponds to absolute DNA accessibility, which can be directly converted to absolute nucleosome occupancy (1-accessibility). Alternatively, DNA methyltransferases (DNMTs) also do not methylate within NCPs and were used to measure accessibility in single-locus studies (MAPit) (Jessen et al., 2004; Kilgore et al., 2007; Small et al., 2014) and even on genome scale in human cells (NOMe-seq) (Kelly et al., 2012). However, all genome-wide studies so far failed to map absolute nucleosome occupancy, as they did not fulfill all of the following prerequisites: frozen nucleosome dynamics (MAPit, *in vivo* in the presence of nucleosome-mobilizing remodelers), saturation of methylation reaction (NOMe-seq, with inappropriate DNMT titration to match unreliable MNase-seq results) and sufficient sequencing coverage (NOMe-seq). Nonetheless, RE accessibility and methylation footprinting are in principle suitable to determine absolute nucleosome occupancies.

Aims of this thesis

In this thesis, I investigated two fundamental aspects of chromatin organization: the determination of genome-wide absolute nucleosome occupancy and the mechanism of nucleosome positioning by ATP dependent remodeling enzymes. The latter aspect was subdivided into nucleosome positioning by remodeler-intrinsic DNA sequence reading versus generation of phased nucleosomal arrays at barriers.

To obtain a genome-wide absolute nucleosome occupancy map, we probed the accessibility of *S. cerevisiae* chromatin with restriction enzymes and DNA methyltransferases. To freeze nucleosome dynamics, we fixed the nucleosomes, and probed accessibility under saturation conditions after titrating the RE or DNMT reactions.

Although it is known that chromatin remodelers with spacing activity are mainly responsible for nucleosomal array generation, it is still unclear what turns sliding activity into spacing activity. Spacing activity is most likely not encoded in the catalytic subunit itself, because it is very conserved and even the same catalytic subunit can lead to different remodeling activities. For example, the ATPase Isw1 mediates spacing in ISW1a but only sliding in ISW1b.
To solve this question, we first focused on the INO80 remodeler, as we had access to recombinant INO80 complexes and structure-based mutant versions through the collaboration with the Hopfner group. We focused on DNA contacts of the Arp8 module and Nhp10 module of the INO80 remodeler, which were suggested to be involved in modulating INO80's sliding activity. We tested several Arp8- and Nhp10-mutants in our genome-wide remodeling assay, which is based on the *in vitro* reconstitution with pure proteins (Krietenstein et al., 2016), monitoring genome-wide remodeling activity on SGD chromatin assembled with a whole-genome plasmid library and *Drosophila melanogaster* embryo histone octamers.

Moreover, we were interested in the DNA sequence read-out mechanism of INO80. To this end, we performed principle component analysis on mononucleosomal fragments obtained after remodeling of SGD chromatin by wild type and mutant INO80 complexes at low nucleosome density, which should reflect the corresponding DNA sequence preferences without much influence by neighboring nucleosomes. Bioinformatic analyses of these sequences should reveal which DNA features are read by INO80 and which INO80 module is involved in such DNA read-out.

Finally, we refined this approach by altering the nucleosome density to investigate remodeler-mediated versus density-driven packing-mechanisms and included all yeast remodelers with spacing activity in combination with Reb1 to investigate spacing mechanisms. We quantified the DNA lengths between nucleosomes, and between Reb1 and nucleosomes. To gain information about INO80's ruler, we tested all Arp8 module- and Nhp10-module mutants. Further, we replaced the underlying DNA sequence with prokaryotic DNA to study the effects of lacking dinucleotide periodicity on spacing activity and array generation.

Abbreviations

Abf1	**A**RS (autonomous replication sequence) **b**inding **f**actor 1
ATAC-seq	**A**ssay for **T**ransposase-**A**ccessible **C**hromatin **seq**uencing
ATP	**A**denosine **T**ri**P**hosphat
CHD	**C**hromodomain **H**elicase **D**NA-binding protein
Cse4	**C**hromosome **se**gregation mutant 4
cryo-EM	**cryo**-**E**lectron **M**icroscopy
DNA	**D**eoxyribo**N**ucleic **A**cid
DNMT	**DN**A **M**ethyl**T**ransferases
DSB	**D**ouble **S**trand **B**reak
FACT	**Fa**cilitates **C**hromatin **T**ranscription
FAIRE-seq	**F**ormaldehyde-**A**ssisted **I**solation of **R**egulatory **E**lements-**seq**uencing
GAL	**GAL**actose metabolism
GRF	**G**eneral **R**egulatory **F**actor
Htz1	**H**istone **t**wo A **z1**
INO80	**INO**sitol requiring
Ioc2/3/4	**I**SWI **o**ne **c**omplex protein 2/3/4
ISWI	**I**mitation **SWI**tch
Itc1	**I**SWI **t**wo **c**omplex protein 1
MAPit	**M**ethyltransferase **A**ccessibility **P**rotocol for **i**ndividual **t**emplates
MNase-ChIP-seq	**M**icrococcal **N**ucle**ase**-**Ch**romatin **I**mmuno**P**recipitation-**seq**uencing
MNase-seq	**M**icrococcal **N**ucle**ase**-**seq**uencing
NDR/NFR	**N**ucleosome-**D**epleted/-**F**ree **R**egion
NOMe-seq	**N**ucleosome **O**ccupancy and **M**ethylome **seq**uencing
PAF	RNA **P**olymerase II-**A**ssociated **F**actor
PHO5	**PHO**sphate metabolism
PIC	**P**re**I**nitiation **C**omplex
SGD	**S**alt **G**radient **D**ialysis
SHL	**S**uper **H**elical **L**ocation
Spt5	**Sp**pressor of **t**y insertion mutant 5
Swi/Snf	**Swi**tch/**S**ucrose-**N**on-**F**ermenting
SWR1	**SW**i/Snf-**R**elated 1
RE	**R**estriction **E**nzyme
Reb1	**R**NA polymerase I **e**nhancer **b**inding protein 1
RecA	**Rec**ombinase **A** (identified in *E. coli*, Rad51 homolog)
RSC	**R**emodels the **S**tructure of **C**hromatin
TSS	**T**ranscription **S**tart **S**ite
WCE	**W**hole **C**ell **E**xtract

Bibliography

Almer, A., Rudolph, H., Hinnen, A., Hörz, W., 1986. Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements. EMBO J. 5, 2689–2696. https://doi.org/10.1002/j.1460-2075.1986.tb04552.x

Anderson, J.D., Widom, J., 2001. Poly(dA-dT) Promoter Elements Increase the Equilibrium Accessibility of Nucleosomal DNA Target Sites. Mol. Cell. Biol. 21, 3830–3839. https://doi.org/10.1128/MCB.21.11.3830-3839.2001

Badis, G., Chan, E.T., van Bakel, H., Pena-Castillo, L., Tillo, D., Tsui, K., Carlson, C.D., Gossett, A.J., Hasinoff, M.J., Warren, C.L., Gebbia, M., Talukder, S., Yang, A., Mnaimneh, S., Terterov, D., Coburn, D., Li Yeo, A., Yeo, Z.X., Clarke, N.D., Lieb, J.D., Ansari, A.Z., Nislow, C., Hughes, T.R., 2008. A Library of Yeast Transcription Factor Motifs Reveals a Widespread Function for Rsc3 in Targeting Nucleosome Exclusion at Promoters. Mol. Cell 32, 878–887. https://doi.org/10.1016/j.molcel.2008.11.020

Bai, L., Morozov, A.V., 2010. Gene regulation by nucleosome positioning. Trends Genet. 26, 476–483. https://doi.org/10.1016/j.tig.2010.08.003

Barbaric, S., Fascher, K.-D., Hoörz, W., 1992. Activation of the weakly regulated PHO8 promoter in S.cerevisiae: chromatin transition and binding sites for the positive regulatory protein PHO4. Nucleic Acids Res. 20, 1031–1038. https://doi.org/10.1093/nar/20.5.1031

Bean, A., Weiner, A., Hughes, A., Itskovits, E., Friedman, N., Rando, O.J., 2016. Genome-wide histone modification patterns in Kluyveromyces Lactis reveal evolutionary adaptation of a heterochromatin-associated mark. bioRxiv 039776. https://doi.org/10.1101/039776

Bell, O., Tiwari, V.K., Thomä, N.H., Schübeler, D., 2011. Determinants and dynamics of genome accessibility. Nat. Rev. Genet. 12, 554–564. https://doi.org/10.1038/nrg3017

Bettecken, T., Trifonov, E.N., 2009. Repertoires of the Nucleosome-Positioning Dinucleotides. PLoS ONE 4. https://doi.org/10.1371/journal.pone.0007654

Bi, X., 2014. Heterochromatin structure: Lessons from the budding yeast. IUBMB Life 66, 657–666. https://doi.org/10.1002/iub.1322

Brahma, S., Udugama, M.I., Kim, J., Hada, A., Bhardwaj, S.K., Hailu, S.G., Lee, T.-H., Bartholomew, B., 2017. INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers. Nat. Commun. 8, 1–12. https://doi.org/10.1038/ncomms15616

Brogaard, K., Xi, L., Wang, J.-P., Widom, J., 2012. A map of nucleosome positions in yeast at base-pair resolution. Nature 486, 496–501. https://doi.org/10.1038/nature11142

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., Greenleaf, W.J., 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods 10, 1213–1218. https://doi.org/10.1038/nmeth.2688

Cairns, B.R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B., Kornberg, R.D., 1996. RSC, an Essential, Abundant Chromatin-Remodeling Complex. Cell 87, 1249–1260. https://doi.org/10.1016/S0092-8674(00)81820-6

Chen, Y., Rai, R., Zhou, Z.-R., Kanoh, J., Ribeyre, C., Yang, Y., Zheng, H., Damay, P., Wang, F., Tsujii, H., Hiraoka, Y., Shore, D., Hu, H.-Y., Chang, S., Lei, M., 2011. A conserved motif within RAP1 has diversified roles in telomere protection and regulation in different organisms. Nat. Struct. Mol. Biol. 18, 213–221. https://doi.org/10.1038/nsmb.1974

Chereji, R.V., Bryson, T.D., Henikoff, S., 2019. Quantitative MNase-seq accurately maps nucleosome occupancy levels. Genome Biol. 20, 1–18. https://doi.org/10.1186/s13059-019-1815-z

Chereji, R.V., Ocampo, J., Clark, D.J., 2017. MNase-Sensitive Complexes in Yeast: Nucleosomes and Non-histone Barriers. Mol. Cell 65, 565-577.e3. https://doi.org/10.1016/j.molcel.2016.12.009

Chereji, R.V., Ramachandran, S., Bryson, T.D., Henikoff, S., 2018. Precise genome-wide mapping of single nucleosomes and linkers in vivo. Genome Biol. 19, 19. https://doi.org/10.1186/s13059-018-1398-0

Clapier, C.R., Cairns, B.R., 2009. The Biology of Chromatin Remodeling Complexes. Annu. Rev. Biochem. 78, 273–304. https://doi.org/10.1146/annurev.biochem.77.062706.153223

Clark, D.J., Kimura, T., 1990. Electrostatic mechanism of chromatin folding. J. Mol. Biol. 211, 883–896. https://doi.org/10.1016/0022-2836(90)90081-V

Clark, R.J., Felsenfeld, G., 1971. Structure of Chromatin. Nature. New Biol. 229, 101–106. https://doi.org/10.1038/newbio229101a0

Colin, J., Candelli, T., Porrua, O., Boulay, J., Zhu, C., Lacroute, F., Steinmetz, L.M., Libri, D., 2014. Roadblock Termination by Reb1p Restricts Cryptic and Readthrough Transcription. Mol. Cell 56, 667–680. https://doi.org/10.1016/j.molcel.2014.10.026

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., Richmond, T.J., 2002. Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9Å Resolution††We dedicate this paper to the memory of Max Perutz who was particularly inspirational and supportive to T.J.R. in the early stages of this study. J. Mol. Biol. 319, 1097–1113. https://doi.org/10.1016/S0022-2836(02)00386-8

David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C.J., Bofkin, L., Jones, T., Davis, R.W., Steinmetz, L.M., 2006. A high-resolution map of transcription in the yeast genome. Proc. Natl. Acad. Sci. 103, 5320–5325. https://doi.org/10.1073/pnas.0601091103

Dekker, J., 2008. Mapping in Vivo Chromatin Interactions in Yeast Suggests an Extended Chromatin Fiber with Regional Variation in Compaction. J. Biol. Chem. 283, 34532–34540. https://doi.org/10.1074/jbc.M806479200

Dingwall, C., Lomonossoff, G.P., Laskey, R.A., 1981. High sequence specificity of micrococcal nuclease. Nucleic Acids Res. 9, 2659–2673.

Eustermann, S., Schall, K., Kostrewa, D., Lakomek, K., Strauss, M., Moldt, M., Hopfner, K.-P., 2018. Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. Nature 556, 386–390. https://doi.org/10.1038/s41586-018-0029-y

Fascher, K.D., Schmitz, J., Hörz, W., 1990. Role of trans-activating proteins in the generation of active chromatin at the PHO5 promoter in S. cerevisiae. EMBO J. 9, 2523–2528. https://doi.org/10.1002/j.1460-2075.1990.tb07432.x

Finch, J.T., Klug, A., 1976. Solenoidal model for superstructure in chromatin. Proc. Natl. Acad. Sci. U. S. A. 73, 1897–1901.

Flaus, A., Martin, D.M.A., Barton, G.J., Owen-Hughes, T., 2006. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. Nucleic Acids Res. 34, 2887–2905. https://doi.org/10.1093/nar/gkl295

Flemming, W., 1882. Zellsubstanz, Kern und Zelltheilung - Google Scholar [WWW Document]. URL https://scholar.google.com/scholar_lookup?title=Zellsubstanz%2C%20Kern%20und%20Zelltheilung&publication_year=1882&author=Flemming%2CW (accessed 10.30.19).

Fourel, G., Miyake, T., Defossez, P.-A., Li, R., Gilson, É., 2002. General Regulatory Factors (GRFs) as Genome Partitioners. J. Biol. Chem. 277, 41736–41743. https://doi.org/10.1074/jbc.M202578200

Fussner, E., Ching, R.W., Bazett-Jones, D.P., 2011. Living without 30 nm chromatin fibers. Trends Biochem. Sci. 36, 1–6. https://doi.org/10.1016/j.tibs.2010.09.002

Garcia-Saez, I., Menoni, H., Boopathi, R., Shukla, M.S., Soueidan, L., Noirclerc-Savoye, M., Le Roy, A., Skoufias, D.A., Bednar, J., Hamiche, A., Angelov, D., Petosa, C., Dimitrov, S., 2018. Structure of an H1-Bound 6-Nucleosome Array Reveals an Untwisted Two-Start Chromatin Fiber Conformation. Mol. Cell 72, 902-915.e7. https://doi.org/10.1016/j.molcel.2018.09.027

Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., Weissman, J.S., 2003. Global analysis of protein expression in yeast. Nature 425, 737–741. https://doi.org/10.1038/nature02046

Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R., Lieb, J.D., 2007. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res. 17, 877–885. https://doi.org/10.1101/gr.5533506

Gkikopoulos, T., Schofield, P., Singh, V., Pinskaya, M., Mellor, J., Smolle, M., Workman, J.L., Barton, G.J., Owen-Hughes, T., 2011. A Role for Snf2-Related Nucleosome-Spacing Enzymes in Genome-Wide Nucleosome Organization. Science 333, 1758–1760. https://doi.org/10.1126/science.1206097

Gossett, A.J., Lieb, J.D., 2012. In vivo effects of histone H3 depletion on nucleosome occupancy and position in Saccharomyces cerevisiae. PLoS Genet. 8, e1002771. https://doi.org/10.1371/journal.pgen.1002771

Grunstein, M., 1997. Molecular model for telomeric heterochromatin in yeast. Curr. Opin. Cell Biol. 9, 383–387. https://doi.org/10.1016/S0955-0674(97)80011-7

Han, M., Grunstein, M., 1988. Nucleosome loss activates yeast downstream promoters in vivo. Cell 55, 1137–1145. https://doi.org/10.1016/0092-8674(88)90258-9

Hansen, J.C., 2012. Human mitotic chromosome structure: what happened to the 30-nm fibre? EMBO J. 31, 1621–1623. https://doi.org/10.1038/emboj.2012.66
Hartley, P.D., Madhani, H.D., 2009. Mechanisms that Specify Promoter Nucleosome Location and Identity. Cell 137, 445–458. https://doi.org/10.1016/j.cell.2009.02.043
Hewish, D.R., Burgoyne, L.A., 1973. Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. Biochem. Biophys. Res. Commun. 52, 504–510. https://doi.org/10.1016/0006-291X(73)90740-7
Holde, K.E. van, 1989. Chromatin, Springer Series in Molecular and Cell Biology. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4612-3490-6
Hörz, W., Altenburger, W., 1981. Sequence specific cleavage of DNA by micrococcal nuclease. Nucleic Acids Res. 9, 2643–2658. https://doi.org/10.1093/nar/9.12.2643
Iyer, V., Struhl, K., 1995. Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. EMBO J. 14, 2570–2579. https://doi.org/10.1002/j.1460-2075.1995.tb07255.x
Jay C. Vary, J., Gangaraju, V.K., Qin, J., Landel, C.C., Kooperberg, C., Bartholomew, B., Tsukiyama, T., 2003. Yeast Isw1p Forms Two Separable Complexes In Vivo. Mol. Cell. Biol. 23, 80–91. https://doi.org/10.1128/MCB.23.1.80-91.2003
Jessen, W.J., Dhasarathy, A., Hoose, S.A., Carvin, C.D., Risinger, A.L., Kladde, M.P., 2004. Mapping chromatin structure in vivo using DNA methyltransferases. Methods, In vitro analysis of chromatin structure in model systems 33, 68–80. https://doi.org/10.1016/j.ymeth.2003.10.025
Jiang, C., Pugh, B.F., 2009. Nucleosome positioning and gene regulation: advances through genomics. Nat. Rev. Genet. 10, 161–172. https://doi.org/10.1038/nrg2522
Joti, Y., Hikima, T., Nishino, Y., Kamada, F., Hihara, S., Takata, H., Ishikawa, T., Maeshima, K., 2012. Chromosomes without a 30-nm chromatin fiber. Nucleus 3, 404–410. https://doi.org/10.4161/nucl.21222
Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., Segal, E., 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. Nature 458, 362–366. https://doi.org/10.1038/nature07667
Kelly, T.K., Liu, Y., Lay, F.D., Liang, G., Berman, B.P., Jones, P.A., 2012. Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. Genome Res. 22, 2497–2506. https://doi.org/10.1101/gr.143008.112
Kilgore, J.A., Hoose, S.A., Gustafson, T.L., Porter, W., Kladde, M.P., 2007. Single-molecule and population probing of chromatin structure using DNA methyltransferases. Methods, Methods Related to the Structure and Function of Eukaryotic Chromatin 41, 320–332. https://doi.org/10.1016/j.ymeth.2006.08.008
Knezetic, J.A., Luse, D.S., 1986. The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. Cell 45, 95–104. https://doi.org/10.1016/0092-8674(86)90541-6
Kornberg, R.D., 1974. Chromatin Structure: A Repeating Unit of Histones and DNA. Science 184, 868–871. https://doi.org/10.1126/science.184.4139.868
Kornberg, R.D., Stryer, L., 1988. Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism. Nucleic Acids Res. 16, 6677–6690.
Kornberg, R.D., Thomas, J.O., 1974. Chromatin Structure: Oligomers of the Histones. Science 184, 865–868. https://doi.org/10.1126/science.184.4139.865
Kowalski, A., Pałyga, J., 2011. Chromatin compaction in terminally differentiated avian blood cells: the role of linker histone H5 and non-histone protein MENT. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. 19, 579–590. https://doi.org/10.1007/s10577-011-9218-3
Krietenstein, N., Wal, M., Watanabe, S., Park, B., Peterson, C.L., Pugh, B.F., Korber, P., 2016. Genomic Nucleosome Organization Reconstituted with Pure Proteins. Cell 167, 709-721.e12. https://doi.org/10.1016/j.cell.2016.09.045
Krietenstein, N., Wippo, C.J., Lieleg, C., Korber, P., 2012. Genome-Wide In Vitro Reconstitution of Yeast Chromatin with In Vivo-Like Nucleosome Positioning, in: Methods in Enzymology. Elsevier, pp. 205–232.

Kubik, S., Bruzzone, M.J., Challal, D., Dreos, R., Mattarocci, S., Bucher, P., Libri, D., Shore, D., 2019. Opposing chromatin remodelers control transcription initiation frequency and start site selection. Nat. Struct. Mol. Biol. 26, 744–754. https://doi.org/10.1038/s41594-019-0273-3

Lai, W.K.M., Pugh, B.F., 2017. Understanding nucleosome dynamics and their links to gene expression and DNA replication. Nat. Rev. Mol. Cell Biol. 18, 548–562. https://doi.org/10.1038/nrm.2017.47

Lantermann, A.B., Straub, T., Strålfors, A., Yuan, G.-C., Ekwall, K., Korber, P., 2010. Schizosaccharomyces pombe genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of Saccharomyces cerevisiae. Nat. Struct. Mol. Biol. 17, 251–257. https://doi.org/10.1038/nsmb.1741

Lawrence, M.B.D., Coutin, N., Choi, J.K., Martin, B.J.E., Irwin, N.A.T., Young, B., Loewen, C., Howe, L.J., 2017. Histone Acetylation, Not Stoichiometry, Regulates Linker Histone Binding in Saccharomyces cerevisiae. Genetics 207, 347–355. https://doi.org/10.1534/genetics.117.1132

Lee, W., Tillo, D., Bray, N., Morse, R.H., Davis, R.W., Hughes, T.R., Nislow, C., 2007. A high-resolution atlas of nucleosome occupancy in yeast. Nat. Genet. 39, 1235–1244. https://doi.org/10.1038/ng2117

Li, M., Hada, A., Sen, P., Olufemi, L., Hall, M.A., Smith, B.Y., Forth, S., McKnight, J.N., Patel, A., Bowman, G.D., Bartholomew, B., Wang, M.D., 2015. Dynamic regulation of transcription factors by nucleosome remodeling. eLife 4, e06249. https://doi.org/10.7554/eLife.06249

Lieleg, C., Ketterer, P., Nuebler, J., Ludwigsen, J., Gerland, U., Dietz, H., Mueller-Planitz, F., Korber, P., 2015a. Nucleosome Spacing Generated by ISWI and CHD1 Remodelers Is Constant Regardless of Nucleosome Density. Mol. Cell. Biol. 35, 1588–1605. https://doi.org/10.1128/MCB.01070-14

Lieleg, C., Krietenstein, N., Walker, M., Korber, P., 2015b. Nucleosome positioning in yeasts: methods, maps, and mechanisms. Chromosoma 124, 131–151. https://doi.org/10.1007/s00412-014-0501-x

Lohr, D., 1997. Nucleosome Transactions on the Promoters of the YeastGAL and PHO Genes. J. Biol. Chem. 272, 26795–26798. https://doi.org/10.1074/jbc.272.43.26795

Lorch, Y., LaPointe, J.W., Kornberg, R.D., 1987. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell 49, 203–210. https://doi.org/10.1016/0092-8674(87)90561-7

Lorch, Y., Maier-Davis, B., Kornberg, R.D., 2014. Role of DNA sequence in chromatin remodeling and the formation of nucleosome-free regions. Genes Dev. 28, 2492–2497. https://doi.org/10.1101/gad.250704.114

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J., 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389, 251–260. https://doi.org/10.1038/38444

Maeshima, K., Hihara, S., Eltsov, M., 2010. Chromatin structure: does the 30-nm fibre exist in vivo? Curr. Opin. Cell Biol., Nucleus and gene expression 22, 291–297. https://doi.org/10.1016/j.ceb.2010.03.001

Maeshima, K., Ide, S., Babokhov, M., 2019. Dynamic chromatin organization without the 30-nm fiber. Curr. Opin. Cell Biol., Cell Nucleus 58, 95–104. https://doi.org/10.1016/j.ceb.2019.02.003

Maeshima, K., Ide, S., Hibino, K., Sasai, M., 2016. Liquid-like behavior of chromatin. Curr. Opin. Genet. Dev., Genome architecture and expression 37, 36–45. https://doi.org/10.1016/j.gde.2015.11.006

Maeshima, K., Imai, R., Tamura, S., Nozaki, T., 2014. Chromatin as dynamic 10-nm fibers. Chromosoma 123, 225–237. https://doi.org/10.1007/s00412-014-0460-2

Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., Pugh, B.F., 2008a. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. Genome Res. 18, 1073–1083. https://doi.org/10.1101/gr.078261.108

Mavrich, T.N., Jiang, C., Ioshikhes, I.P., Li, X., Venters, B.J., Zanton, S.J., Tomsho, L.P., Qi, J., Glaser, R.L., Schuster, S.C., Gilmour, D.S., Albert, I., Pugh, B.F., 2008b. Nucleosome organization in the Drosophila genome. Nature 453, 358–362. https://doi.org/10.1038/nature06929

McDowall, A. w., Smith, J. m., Dubochet, J., 1986. Cryo-electron microscopy of vitrified chromosomes in situ. EMBO J. 5, 1395–1402. https://doi.org/10.1002/j.1460-2075.1986.tb04373.x

McGinty, R.K., Tan, S., 2015. Nucleosome Structure and Function. Chem. Rev. 115, 2255–2273. https://doi.org/10.1021/cr500373h
Meluh, P.B., Yang, P., Glowczewski, L., Koshland, D., Smith, M.M., 1998. Cse4p Is a Component of the Core Centromere of Saccharomyces cerevisiae. Cell 94, 607–613. https://doi.org/10.1016/S0092-8674(00)81602-5
Mirny, L.A., Imakaev, M., Abdennur, N., 2019. Two major mechanisms of chromosome organization. Curr. Opin. Cell Biol. 58, 142–152. https://doi.org/10.1016/j.ceb.2019.05.001
Mizuguchi, G., Shen, X., Landry, J., Wu, W.-H., Sen, S., Wu, C., 2004. ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex. Science 303, 343–348. https://doi.org/10.1126/science.1090701
Möbius, W., Gerland, U., 2010. Quantitative Test of the Barrier Nucleosome Model for Statistical Positioning of Nucleosomes Up- and Downstream of Transcription Start Sites. PLOS Comput. Biol. 6, e1000891. https://doi.org/10.1371/journal.pcbi.1000891
Nishino, Y., Eltsov, M., Joti, Y., Ito, K., Takata, H., Takahashi, Y., Hihara, S., Frangakis, A.S., Imamoto, N., Ishikawa, T., Maeshima, K., 2012. Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. EMBO J. 31, 1644–1653. https://doi.org/10.1038/emboj.2012.35
Noll, M., 1974. Subunit structure of chromatin. Nature 251, 249–251. https://doi.org/10.1038/251249a0
Ocampo, J., Chereji, R.V., Eriksson, P.R., Clark, D.J., 2016. The ISW1 and CHD1 ATP-dependent chromatin remodelers compete to set nucleosome spacing in vivo. Nucleic Acids Res. 44, 4625–4635. https://doi.org/10.1093/nar/gkw068
Olins, A.L., Olins, D.E., 1974. Spheroid Chromatin Units (ν Bodies). Science 183, 330–332. https://doi.org/10.1126/science.183.4122.330
Olins, D.E., Olins, A.L., 2003. Chromatin history: our view from the bridge. Nat. Rev. Mol. Cell Biol. 4, 809–814. https://doi.org/10.1038/nrm1225
Oudet, P., Gross-Bellard, M., Chambon, P., 1975. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell 4, 281–300. https://doi.org/10.1016/0092-8674(75)90149-X
Papamichos-Chronakis, M., Watanabe, S., Rando, O.J., Peterson, C.L., 2011. Global Regulation of H2A.Z Localization by the INO80 Chromatin-Remodeling Enzyme Is Essential for Genome Integrity. Cell 144, 200–213. https://doi.org/10.1016/j.cell.2010.12.021
Parnell, T.J., Huff, J.T., Cairns, B.R., 2008. RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. EMBO J. 27, 100–110. https://doi.org/10.1038/sj.emboj.7601946
Patterton, H.G., Landel, C.C., Landsman, D., Peterson, C.L., Simpson, R.T., 1998. The Biochemical and Phenotypic Characterization of Hho1p, the Putative Linker Histone H1 of Saccharomyces cerevisiae. J. Biol. Chem. 273, 7268–7276. https://doi.org/10.1074/jbc.273.13.7268
Peterson, C.L., Laniel, M.-A., 2004. Histones and histone modifications. Curr. Biol. 14, R546–R551. https://doi.org/10.1016/j.cub.2004.07.007
Radman-Livaja, M., Rando, O.J., 2010. Nucleosome positioning: How is it established, and why does it matter? Dev. Biol., Special Section: Gene Expression and Development 339, 258–266. https://doi.org/10.1016/j.ydbio.2009.06.012
Reed, S.H., Akiyama, M., Stillman, B., Friedberg, E.C., 1999. Yeast autonomously replicating sequence binding factor is involved in nucleotide excision repair. Genes Dev. 13, 3052–3058. https://doi.org/10.1101/gad.13.23.3052
Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M., Cosma, M.P., 2015. Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo. Cell 160, 1145–1158. https://doi.org/10.1016/j.cell.2015.01.054
Schones, D.E., Cui, K., Cuddapah, S., Roh, T.-Y., Barski, A., Wang, Z., Wei, G., Zhao, K., 2008. Dynamic Regulation of Nucleosome Positioning in the Human Genome. Cell 132, 887–898. https://doi.org/10.1016/j.cell.2008.02.022
Sedat, J., Manuelidis, L., 1978. A Direct Approach to the Structure of Eukaryotic Chromosomes. Cold Spring Harb. Symp. Quant. Biol. 42, 331–350. https://doi.org/10.1101/SQB.1978.042.01.035
Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I.K., Wang, J.-P.Z., Widom, J., 2006. A genomic code for nucleosome positioning. Nature 442, 772–778. https://doi.org/10.1038/nature04979

Shen, X., Mizuguchi, G., Hamiche, A., Wu, C., 2000. A chromatin remodelling complex involved in transcription and DNA processing. Nature 406, 541–544. https://doi.org/10.1038/35020123
Shivaswamy, S., Bhinge, A., Zhao, Y., Jones, S., Hirst, M., Iyer, V.R., 2008. Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. PLoS Biol. 6, e65. https://doi.org/10.1371/journal.pbio.0060065
Simic, R., Lindstrom, D.L., Tran, H.G., Roinick, K.L., Costa, P.J., Johnson, A.D., Hartzog, G.A., Arndt, K.M., 2003. Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. EMBO J. 22, 1846–1856. https://doi.org/10.1093/emboj/cdg179
Simpson, R.T., 1978. Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochemistry 17, 5524–5531. https://doi.org/10.1021/bi00618a030
Small, E.C., Xi, L., Wang, J.-P., Widom, J., Licht, J.D., 2014. Single-cell nucleosome mapping reveals the molecular basis of gene expression heterogeneity. Proc. Natl. Acad. Sci. 111, E2462–E2471. https://doi.org/10.1073/pnas.1400517111
Smith, C.D., Shu, S., Mungall, C.J., Karpen, G.H., 2007. The Release 5.1 Annotation of Drosophila melanogaster Heterochromatin. Science 316, 1586–1591. https://doi.org/10.1126/science.1139815
Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.-M., Zhu, P., Li, G., 2014. Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units. Science 344, 376–380. https://doi.org/10.1126/science.1251413
Stockdale, C., Flaus, A., Ferreira, H., Owen-Hughes, T., 2006. Analysis of Nucleosome Repositioning by Yeast ISWI and Chd1 Chromatin Remodeling Complexes. J. Biol. Chem. 281, 16279–16288. https://doi.org/10.1074/jbc.M600682200
Talbert, P.B., Henikoff, S., 2017. Histone variants on the move: substrates for chromatin dynamics. Nat. Rev. Mol. Cell Biol. 18, 115–126. https://doi.org/10.1038/nrm.2016.148
Telford, D.J., Stewart, B.W., 1989. Micrococcal nuclease: Its specificity and use for chromatin analysis. Int. J. Biochem. 21, 127–138. https://doi.org/10.1016/0020-711X(89)90100-6
Tirosh, I., Barkai, N., 2008. Two strategies for gene regulation by promoter nucleosomes. Genome Res. 18, 1084–1091. https://doi.org/10.1101/gr.076059.108
Tirosh, I., Sigal, N., Barkai, N., 2010. Divergence of nucleosome positioning between two closely related yeast species: genetic basis and functional consequences. Mol. Syst. Biol. 6, 365. https://doi.org/10.1038/msb.2010.20
Udugama, M., Sabri, A., Bartholomew, B., 2011. The INO80 ATP-Dependent Chromatin Remodeling Complex Is a Nucleosome Spacing Factor. Mol. Cell. Biol. 31, 662–673. https://doi.org/10.1128/MCB.01035-10
van Bakel, H., Tsui, K., Gebbia, M., Mnaimneh, S., Hughes, T.R., Nislow, C., 2013. A Compendium of Nucleosome and Transcript Profiles Reveals Determinants of Chromatin Architecture and Transcription. PLoS Genet. 9. https://doi.org/10.1371/journal.pgen.1003479
Wang, F., Ranjan, A., Wei, D., Wu, C., 2016. Comment on "A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme." Science 353, 358–358. https://doi.org/10.1126/science.aad5921
Weiner, A., Hughes, A., Yassour, M., Rando, O.J., and Friedman, N. (2010). High-resolution nucleosome mapping reveals transcription-dependent promoter packaging. Genome Res. *20*, 90–100.
Willhoft, O., Ghoneim, M., Lin, C.-L., Chua, E.Y.D., Wilkinson, M., Chaban, Y., Ayala, R., McCormack, E.A., Ocloo, L., Rueda, D.S., Wigley, D.B., 2018. Structure and dynamics of the yeast SWR1-nucleosome complex. Science 362. https://doi.org/10.1126/science.aat7716
Woodcock, C.L., Frado, L.L., Rattner, J.B., 1984. The higher-order structure of chromatin: evidence for a helical ribbon arrangement. J. Cell Biol. 99, 42–52. https://doi.org/10.1083/jcb.99.1.42
Woodcock, C.L., Skoultchi, A.I., Fan, Y., 2006. Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. Chromosome Res. 14, 17–25. https://doi.org/10.1007/s10577-005-1024-3
Woodcock, C.L.F., Safer, J.P., Stanchfield, J.E., 1976. Structural repeating units in chromatin: I. Evidence for their general occurrence. Exp. Cell Res. 97, 101–110. https://doi.org/10.1016/0014-4827(76)90659-5

Workman, J.L., Roeder, R.G., 1987. Binding of transcription factor TFIID to the major late promoter during in vitro nucleosome assembly potentiates subsequent initiation by RNA polymerase II. Cell 51, 613–622. https://doi.org/10.1016/0092-8674(87)90130-9

Yamada, K., Frouws, T.D., Angst, B., Fitzgerald, D.J., DeLuca, C., Schimmele, K., Sargent, D.F., Richmond, T.J., 2011. Structure and mechanism of the chromatin remodelling factor ISW1a. Nature 472, 448–453. https://doi.org/10.1038/nature09947

Yan, C., Chen, H., Bai, L., 2018. Systematic Study of Nucleosome-Displacing Factors in Budding Yeast. Mol. Cell 71, 294-305.e4. https://doi.org/10.1016/j.molcel.2018.06.017

Yang, J.G., Madrid, T.S., Sevastopoulos, E., Narlikar, G.J., 2006. The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. Nat. Struct. Mol. Biol. 13, 1078–1083. https://doi.org/10.1038/nsmb1170

Zhang, Y., Moqtaderi, Z., Rattner, B.P., Euskirchen, G., Snyder, M., Kadonaga, J.T., Liu, X.S., Struhl, K., 2009. Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. Nat. Struct. Mol. Biol. 16, 847–852. https://doi.org/10.1038/nsmb.1636

Zhang, Z., Wippo, C.J., Wal, M., Ward, E., Korber, P., Pugh, B.F., 2011. A Packing Mechanism for Nucleosome Organization Reconstituted Across a Eukaryotic Genome. Science 332, 977–980. https://doi.org/10.1126/science.1200508

Zhao, Z., Shilatifard, A., 2019. Epigenetic modifications of histones in cancer. Genome Biol. 20, 1–16. https://doi.org/10.1186/s13059-019-1870-5

Zhou, C.Y., Johnson, S.L., Gamarra, N.I., Narlikar, G.J., 2016. Mechanisms of ATP-Dependent Chromatin Remodeling Motors. Annu. Rev. Biophys. 45, 153–181. https://doi.org/10.1146/annurev-biophys-051013-022819

Zhou, C.Y., Johnson, S.L., Lee, L.J., Longhurst, A.D., Beckwith, S.L., Johnson, M.J., Morrison, A.J., Narlikar, G.J., 2018. The Yeast INO80 Complex Operates as a Tunable DNA Length-Sensitive Switch to Regulate Nucleosome Sliding. Mol. Cell 69, 677-688.e9. https://doi.org/10.1016/j.molcel.2018.01.028

Zink, L.-M., Hake, S.B., 2016. Histone variants: nuclear function and disease. Curr. Opin. Genet. Dev., Genome architecture and expression 37, 82–89. https://doi.org/10.1016/j.gde.2015.12.002

CHAPTER 1: ABSOLUTE NUCLEOSOME OCCUPANCY MAP FOR THE *SACCHAROMYCES CEREVISIAE* GENOME

Elisa Oberbeckmann[1,8], Michael Wolff[2,8], Nils Krietenstein[1,3], Mark Heron[4,5], Jessica L. Ellins[6], Andrea Schmid[1], Stefan Krebs[7], Helmut Blum[7], Ulrich Gerland[2] and Philipp Korber[1]

[1]Molecular Biology Division, Biomedical Center, Faculty of Medicine, Ludwig-Maximilians-Universität München, Planegg- Martinsried, Germany; [2]Physik Department, Technische Universität München, Garching, Germany; [3]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts, USA; [4]Quantitative and Computational Biology, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany; [5]Gene Center, Faculty of Chemistry and Pharmacy, Ludwig-Maximilians-Universität München, Munich, Germany; [6]Department of Biochemistry, University of Oxford, Oxford, United Kingdom; [7]Laboratory of Functional Genome Analysis (LAFUGA), Gene Center, Faculty of Chemistry and Pharmacy, Ludwig-Maximilians-Universität München, Munich, Germany; [8]These authors contributed equally to the work.

This chapter was published in Genome Research.
DOI: 10.1101/gr.253419.119

Author contributions to "Absolute nucleosome occupancy map for the Saccharomyces cerevisiae genome"

I developed the ORE- and ODM-seq method in the lab based on the concepts and ideas from P.K., N.K., M.W., M.H and me. I performed all ODM-seq and in vitro reconstitution experiments. Further, I performed all ORE-seq experiments or supervised them, if they were conducted by A.S. or J.L.E. Nanopore and Illumina sequencing libraries were sequenced with L.A.F.U.G.A by S.K. Data was analyzed by M.W. based on initial analysis from M.H. Figures 2 and 4 was partially, and 3 completely plotted and assembled by me based on data analysis from M.W. Figure 1 was plotted by M.W. and assembled by me. Funding acquisition, project administration and supervision was done by H.B., U.G. and P.K. Results were validated by N.K., A.S., J.L.E., S.K., P.K. and me. Original draft was written by P.K. (Introduction, Results and Discussion), and M.W. and me (Methods). Draft was reviewed and edited by M.W., N.K., M.H., U.G., P.K. and me.

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

Method

Absolute nucleosome occupancy map for the *Saccharomyces cerevisiae* genome

Elisa Oberbeckmann,[1,8] Michael Wolff,[2,8] Nils Krietenstein,[1,3] Mark Heron,[4,5] Jessica L. Ellins,[6] Andrea Schmid,[1] Stefan Krebs,[7] Helmut Blum,[7] Ulrich Gerland,[2] and Philipp Korber[1]

[1]*Molecular Biology Division, Biomedical Center, Faculty of Medicine, Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany;* [2]*Physik Department, Technische Universität München, 85748 Garching, Germany;* [3]*Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA;* [4]*Quantitative and Computational Biology, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany;* [5]*Gene Center, Faculty of Chemistry and Pharmacy, Ludwig-Maximilians-Universität München, 81377 Munich, Germany;* [6]*Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, United Kingdom;* [7]*Laboratory of Functional Genome Analysis (LAFUGA), Gene Center, Faculty of Chemistry and Pharmacy, Ludwig-Maximilians-Universität München, 81377 Munich, Germany*

Mapping of nucleosomes, the basic DNA packaging unit in eukaryotes, is fundamental for understanding genome regulation because nucleosomes modulate DNA access by their positioning along the genome. A cell-population nucleosome map requires two observables: nucleosome positions along the DNA ("Where?") and nucleosome occupancies across the population ("In how many cells?"). All available genome-wide nucleosome mapping techniques are yield methods because they score either nucleosomal (e.g., MNase-seq, chemical cleavage-seq) or nonnucleosomal (e.g., ATAC-seq) DNA but lose track of the total DNA population for each genomic region. Therefore, they only provide nucleosome positions and maybe compare relative occupancies between positions, but cannot measure absolute nucleosome occupancy, which is the fraction of all DNA molecules occupied at a given position and time by a nucleosome. Here, we established two orthogonal and thereby cross-validating approaches to measure absolute nucleosome occupancy across the *Saccharomyces cerevisiae* genome via restriction enzymes and DNA methyltransferases. The resulting high-resolution (9-bp) map shows uniform absolute occupancies. Most nucleosome positions are occupied in most cells: 97% of all nucleosomes called by chemical cleavage-seq have a mean absolute occupancy of 90 ± 6% (±SD). Depending on nucleosome position calling procedures, there are 57,000 to 60,000 nucleosomes per yeast cell. The few low absolute occupancy nucleosomes do not correlate with highly transcribed gene bodies, but correlate with increased presence of the nucleosome-evicting chromatin structure remodeling (RSC) complex, and are enriched upstream of highly transcribed or regulated genes. Our work provides a quantitative method and reference frame in absolute terms for future chromatin studies.

[Supplemental material is available for this article.]

It makes a fundamental difference whether a measurement yields values on a relative or absolute scale. In molecular biology, it is usually much easier to obtain relative values by comparing samples to references than to measure in absolute terms. For example, protein concentration is readily compared between cells by western blotting. However, the absolute number of protein molecules per cell is more difficult to obtain. Molecular biology generally suffers from the scarcity of quantitative data on absolute scales, which hinders a deeper understanding, modeling, and theoretical description of mechanisms and systems features.

Here, we amend such lack of absolute values for the basic packaging unit of eukaryotic genomes: the nucleosome. It is defined as a core of 147 bp of DNA spooled in 1.7 turns around a histone protein octamer (Luger et al. 1997) plus variable lengths of flanking linker DNA. For brevity and according to common usage, "nucleosome" refers to the nucleosome core in the following. The close interactions between DNA and histones in nucleosomes inhibit DNA access for many factors and thereby constitute an important level of regulation for all DNA-dependent processes, like transcription or DNA repair (Venkatesh and Workman 2015; Seeber et al. 2018). Therefore, the mapping of nucleosomes, the dynamics of their positioning and composition, as well as their roles in genome regulation are of paramount interest.

Several techniques map nucleosome positions along a genome (Jiang and Pugh 2009; Meyer and Liu 2014; Lieleg et al. 2015b). The most common tool is micrococcal nuclease (MNase), which digests nonnucleosomal DNA faster than nucleosomal DNA and yields mononucleosomes at a properly limited digestion degree (Rill and Van Holde 1973; Noll 1974). High-throughput sequencing of mononucleosomal DNA maps histone octamer DNA footprints (MNase-seq) (Albert et al. 2007). Because other DNA-bound factors may also inhibit MNase digestion and yield mononucleosome-sized DNA fragments (Chereji et al. 2017), additional criteria can ensure nucleosome specificity, for example, mononucleosome selection by anti-histone

[8]**These authors contributed equally to the work.**
Corresponding author: pkorber@lmu.de
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.253419.119.

© 2019 Oberbeckmann et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

29:1996–2009 Published by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/19; www.genome.org

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

immunoprecipitation (MNase-anti-histone-chromatin immunoprecipitation-seq [MNnase-anti-histone-ChIP-seq]) (Albert et al. 2007; Wal and Pugh 2012). Alternatively, chemical cleavage (Flaus et al. 1996) is an MNase-independent and histone-specific method to map nucleosomes, especially its most recent version (Chereji et al. 2018). Cysteine residues are introduced in histones and coupled to copper-chelating phenanthroline so that incubation with hydrogen peroxide generates hydroxyl radicals that cleave defined DNA sites in the nucleosome. High-throughput sequencing of the resulting DNA fragments yields very precise genome-wide nucleosome maps (Brogaard et al. 2012; Moyle-Heyrman et al. 2013; Chereji et al. 2018). The flip side to measuring nucleosome positions is mapping linker DNA between nucleosome cores and wider nucleosome-free regions (NFRs), for example, by using a hyperactive transposase that inserts sequencing adapters into nucleosome-free DNA (assay for transposase-accessible chromatin [ATAC-seq]) (Buenrostro et al. 2013).

Although such techniques measure nucleosome positions or nonnucleosomal regions and may provide their relative occupancies, they all cannot measure absolute nucleosome occupancy. This quantity describes the fraction of a molecule/cell population where a certain base pair either is part of any nucleosome (Kaplan et al. 2010) or at a nucleosome center (Zhang et al. 2009; Lieleg et al. 2015b). The nucleosome center is also called dyad because of the pseudo-twofold nucleosome symmetry. In the following, we use the first definition because our approach scores the whole nucleosome footprint and not dyads. This fraction can vary between 0% (no molecule has any nucleosome covering this position) and 100% (all molecules have a nucleosome here). The aforementioned methods cannot measure this fraction because they are yield methods, meaning they only score either the nucleosomal or the nonnucleosomal amount at a time but lose track of the total population. A genomic position may yield more or less, for example, MNase-seq or ATAC-seq signal than another position, which is often interpreted as differential nucleosome occupancy (Jiang and Pugh 2009). However, this only refers to occupancy differences in *relative* terms. It remains unknown to which fraction of the total, that is, to which *absolute* occupancy such signals correspond.

Additionally, MNase-based methods suffer from further problems. First, MNase has some sequence bias (Dingwall et al. 1981; Hörz and Altenburger 1981) because it cleaves AT-rich DNA more readily, even within nucleosomes (Cockell et al. 1983; Caserta et al. 2009; Chereji et al. 2019a). Second, MNase digestion has to be limited, otherwise it also cleaves within nucleosomes. Thus, MNase-based methods do not operate at saturation. Standardization or normalization of digestion degrees is very challenging (Cole et al. 2011; Rizzo et al. 2012), but would be necessary because peak heights depend on the digestion degree (Weiner et al. 2010; DeGennaro et al. 2013; Chereji et al. 2019a). MNase must cut on both sides of the nucleosome to cut out a mononucleosome with a probability that depends on the digestion degree as well as on the sequence and length of the flanking DNA. Accordingly, long and AT-rich NFRs, which are typical for budding yeast promoters (Yuan et al. 2005), are frequently cut already at low digestion degrees leading to highest MNase-seq peaks for their flanking −1 and +1 nucleosomes, but are also the entry way for more efficient nucleosome digestion at higher digestion degrees, then leading to relatively lower flanking nucleosome peaks (Zhang et al. 2009; Stein et al. 2010; Weiner et al. 2010; Givens et al. 2012; DeGennaro et al. 2013; Flores et al. 2014). Only the combination of spike-in normalization controls and digestion degree titration allows quantitative relative, albeit still not absolute occupancy measurements by MNase (Chereji et al. 2019a).

Collectively, and as noted before (Rizzo et al. 2011; Ozonov and van Nimwegen 2013; Quintales et al. 2015), we are still blind to absolute nucleosome occupancies on the genome-scale, despite genome-wide mapping of nucleosome positions with increasing precision over the last 14 years (Yuan et al. 2005; Lieleg et al. 2015b).

Nonetheless, there are absolute nucleosome occupancy measurements at single loci based on the differential accessibility of nucleosomal versus nonnucleosomal DNA for restriction enzymes (REs) and DNA methyltransferases (DNMTs). Both enzymes will only cleave/methylate nonnucleosomal DNA, but not remove nucleosomal DNA. If both DNA types are monitored, the fraction of not cleaved/methylated molecules directly corresponds to the absolute occupancy. RE accessibility measurements by Southern blotting were pioneered, for example, at the budding yeast *PHO5* and *PHO8* promoters (Almer et al. 1986; Barbaric et al. 1992). DNA methylation footprinting was established using prokaryotic DNMTs that methylate either CpG or GpC sites (Jessen et al. 2004; Kilgore et al. 2007; Small et al. 2014).

To measure absolute occupancy, it is crucial that the RE- or DNMT-catalyzed reactions reach saturation, which requires that nucleosome dynamics are frozen. This is mostly true for ex vivo–prepared or in vitro–assembled chromatin under physiological buffer and temperature conditions (Korolev et al. 2007; Zhang et al. 2011), but not in vivo where ATP-dependent nucleosome remodeling enzymes (remodelers) reposition, disassemble, or restructure nucleosomes and generate transient DNA accessibility also within nucleosomes (Bartholomew 2014; Zhou et al. 2016).

Genome-wide methods using REs or DNMTs have not yet provided reliable absolute occupancies. Comparisons of RE accessibilities—NA-seq (Gargiulo et al. 2009) and RED-seq (Chen et al. 2014)—scored only cut fragments, for example, by ligating biotinylated adapters after RE digestion and sequencing only the streptavidin-immunoprecipitated DNA (RED-seq). DNMT approaches, for example, NOMe-seq (Kelly et al. 2012), were insufficient mainly for two reasons. First, genome sequencing coverage is often much too low, especially for metazoans, so that only coarsely grained absolute occupancy values could be discerned. Second, DNA methylation extent was either not saturating or chosen to match MNase-seq results. DNA methylation titration for establishing the NOMe-seq protocol (Kelly et al. 2012) led to a discrepancy at the human *MLH1* locus where the GpC-specific DNMT M.CviPI methylated its cognate site even though MNase-seq gave a "nucleosome" signal here. The authors interpreted that the DNMT methylated within a nucleosome and accordingly chose less extensive methylation conditions, which were subsequently used by many in the field (Krebs et al. 2017; Levo et al. 2017).

Considering the aforementioned limitations of existing and especially MNase-based methods, we argue that a reliable and MNase-independent assessment of genome-wide absolute nucleosome occupancy is needed, and we hereby provide it for the *Saccharomyces cerevisiae* genome.

Results

ORE-seq: genome-wide absolute occupancy measurement by REs

As a first approach to measure genome-wide absolute occupancy at low resolution, we brought classical restriction enzyme accessibility assays (Almer et al. 1986; Gregory et al. 1999) to the genome level (Supplemental Methods) and called this method occupancy

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

measurement via restriction enzymes and high-throughput sequencing (ORE-seq). Chromatin was prepared from logarithmically growing (Supplemental Fig. S1A) wild-type (WT) *S. cerevisiae* and digested with REs at different RE concentrations and for different incubation times to establish saturation (Fig. 1A, left). DNA purified after RE digest was sonicated and Illumina-sequenced in paired-end mode. Each fragment end was scored as generated by RE cut versus by sonication, primarily depending on whether its end is close to an RE site or not. Further, we corrected for fortuitous sonication breaks at RE sites, end resection at RE sites attributed to endogenous exonucleases, and deviations from calibration curves generated by analyzing mixtures of cut and uncut genomic DNA at defined ratios (Supplemental Methods).

The application of ORE-seq to biological chromatin replicates using different REs showed good reproducibility and clear saturation of RE digestions; that is, mean absolute occupancy values for each RE were within five percentage points for samples varying by different RE concentrations or incubation times (Supplemental Fig. S1B). As technical control, we exploited that each 6-bp HindIII site contains a 4-bp AluI site. Absolute occupancies measured at HindIII sites by high concentrations of either AluI or HindIII agreed well within 7% mean absolute difference for each site and 2% difference in mean sample occupancy averaged over all sites (mean absolute occupancy) (Supplemental Fig. S1C).

We selected absolute occupancy data for each RE site according to quality criteria, like saturation of digestion or sequencing coverage (Supplemental Methods), and combined them into a genome-wide ORE-seq map with low average resolution of approximately 870 bp (Fig. 1B).

ODM-seq: genome-wide absolute occupancy measurement by DNMTs

For increased resolution and as an orthogonal method, we established (Supplemental Methods) genome-wide absolute occupancy measurement in chromatin by differential cytosine methylation at position C-5 in CpG or GpC motifs and called it occupancy measurement via DNA methylation and high-throughput sequencing (ODM-seq) (Fig. 1A, right). Comparing DNA methylation of in vitro reconstituted nucleosomes using fly histones with ex vivo prepared yeast chromatin, we found that yeast nucleosomes were inherently unstable during prolonged incubation required for saturation of methylation unless magnesium was added or the chromatin was formaldehyde cross-linked (Supplemental Fig. S2A–E; Supplemental Methods). To avoid endogenous nucleases active in the presence of magnesium, we used cross-linked chromatin for ODM-seq (Supplemental Fig. S3; Supplemental Methods).

Methylated cytosines (5mC) were mostly detected by Illumina sequencing of sonication fragments after treatment with bisulfite (BS-seq), which converts only unmethylated cytosines to uracil and thereby changes the DNA sequence in a tractable way. Enzymatic (EM-seq, New England Biolabs) instead of bisulfite conversion in short or direct readout of 5mC in long DNA fragments by Oxford Nanopore sequencing resulted in equivalent occupancy maps (Fig. 1C). This controls against any systematic errors in our bisulfite sequencing and bioinformatics pipeline. Further, Oxford Nanopore sequencing excels in sequencing very long fragments (>10 kb) but did not yield completely unmethylated long DNA fragments. This argues against a contribution of DNA from unlysed cells, which would not have been accessible to REs or DNMTs and may have systematically distorted absolute occupancy measurements. Therefore, we conclude that known biases of bisulfite sequencing (Darst et al. 2010) did not affect our results.

Comparison of different methods and conditions with regard to absolute occupancy measurements

Now we had absolute occupancy measurements across the yeast genome for two independent methods (ORE-seq and ODM-seq) involving two different conditions (non-cross-linked chromatin in RE buffer with Mg^{2+} vs. cross-linked chromatin in DNMT buffer without Mg^{2+}), five independent enzymes (AluI, BamHI, HindIII, M.SssI [CpG DNMT], M.CviPI [GpC DNMT]), five independent biological replicates (WT1 to WT5), two independent technical replicates for BamHI, and a comparison between purified in vitro chromatin and complex ex vivo chromatin for BamHI and the two DNMTs (Supplemental Fig. S2A). All these independent and partially orthogonal measurements yielded mean absolute occupancy values in the range of 71%–81% (Fig. 1D). Because there was no genome-wide precedent for such values, it was important that this multitude of approaches converged in a similar range and cross-validated each other.

Nonetheless, ORE-seq tended to yield lower mean occupancy values than ODM-seq (Fig. 1B,D) and CpG methylation yielded lower mean occupancy values than GpC methylation (Fig. 1D; Supplemental Fig. S2A,B). These differences were only in some cases significant, for example, GpC DNMT versus AluI or HindIII, and unlikely due to different conditions, methods, or replicates as values for BamHI and CpG DNMT overlapped across these differences (Fig. 1D). We considered that different enzymes may probe the same chromatin in different ways. Mean absolute occupancy values across the genome are influenced by the different distribution of cognate sites for each enzyme in nucleosomes versus nonnucleosomal regions (Supplemental Fig. S4A). To correct for such different site distributions, we used nucleosome positions mapped by chemical cleavage (Chereji et al. 2018) and plotted absolute occupancy values for each enzyme as averaged over 20-bp bins around the dyads of all called nucleosomes (Fig. 1E). Here, we also included AluI data from an independent approach called qDA-seq (Chereji et al. 2019b) that is conceptually equivalent to our ORE-seq and became available during the revision phase of our manuscript. This plot showed how each enzyme differently measured occupancy in and around nucleosomes, including linker regions, and the DNA at the entry and exit sites of the histone octamer that is known to transiently unwind (Polach and Widom 1995). For most enzyme pairwise comparisons, the differences for maximal and minimal occupancies, that is, at nucleosome dyads and in linkers, respectively, were within the error. Only GpC DNMT and our AluI data were just significantly different; that is, error bars almost touched. CpG DNMT, our AluI and HindIII, showed a larger min-max difference than BamHI, qDA-AluI, and GpC DNMT, which may mean that the former enzymes can access linker regions and entry/exit DNA more efficiently. However, these differences are within the mean standard deviation and barely significant. The average AluI occupancy values from us versus from the Clark group (Chereji et al. 2019b) agreed within 1% at dyads and differed by 9% in linkers, with almost touching error bars for the latter. When the data was gene averaged and +1 nucleosome-aligned, the qDA-AluI data showed a clear trend toward higher occupancies in linkers and NFRs (Supplemental Fig. S4B). This may be a result of different technical procedures or different biological conditions. For example, the Clark group (Chereji et al. 2019b) used a different strain background (W303), different media (synthetic complete),

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

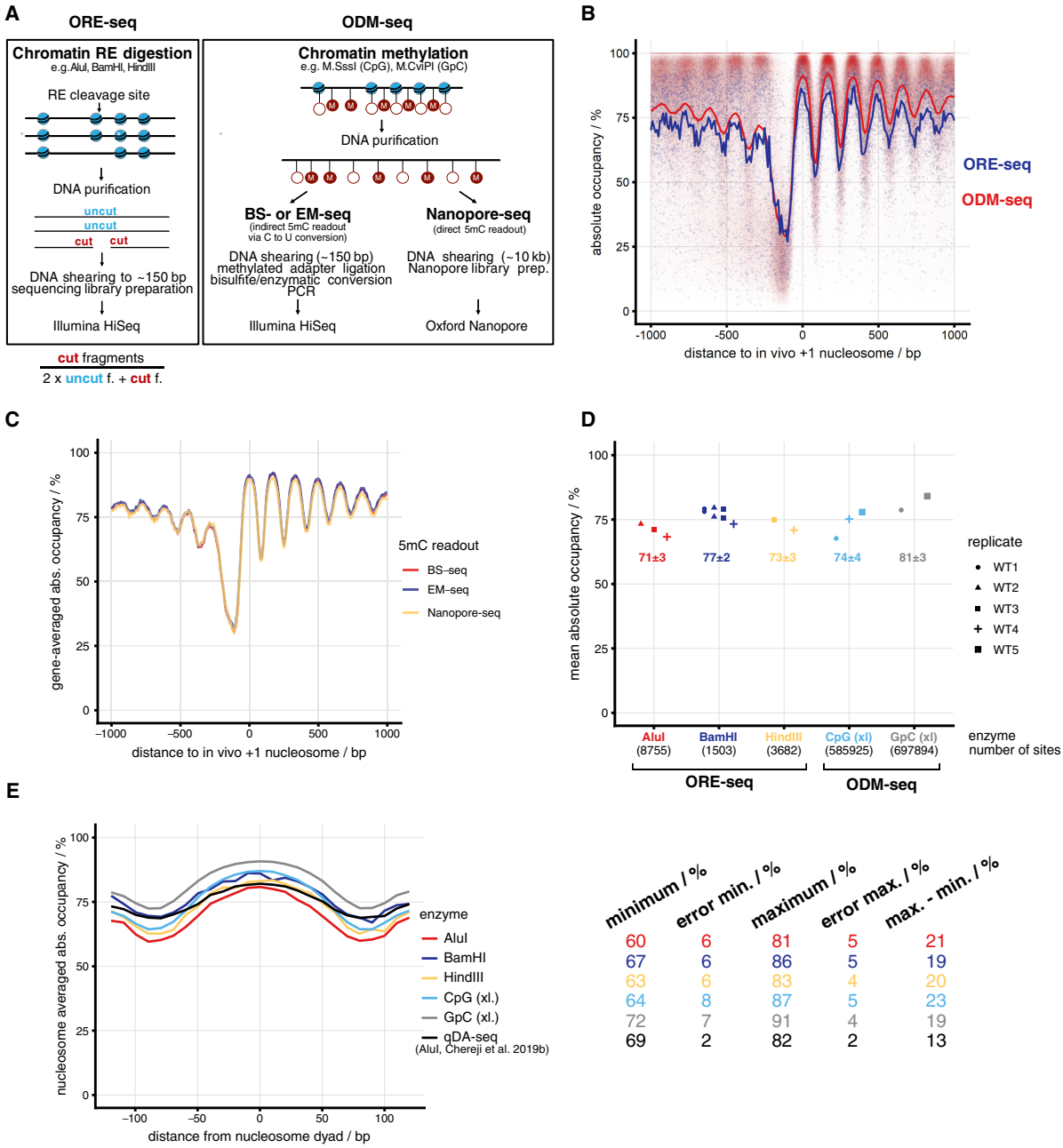


Figure 1. Genome-wide absolute occupancy measurement by restriction enzymes and DNA methyltransferases. (*A*) Method overview; for details, see text. Lollipops stand for DNA methylation sites: (open) unmethylated; (red fill "M") methylated; (f.) fragments. (*B*) Composite plot of absolute occupancy ORE-seq and ODM-seq data averaged over all included samples and aligned at in vivo +1 nucleosome positions. Each dot represents the value of one genomic site, and the lines show the 10-bp bin mean occupancy of aligned sites. (*C*) As in *B*, but for WT5 replicate and the different 5mC readouts stated in *A*. (*D*) Absolute occupancy averaged over all sites (mean absolute occupancy) obtained by ORE-seq (Supplemental Table S3) or ODM-seq (Supplemental Table S4) for the indicated enzymes and biological replicates (WT1 to WT5) at saturation conditions. The number of sites implemented for each enzyme is indicated. (xl.) in vivo formaldehyde cross-linked. (*E*) Absolute occupancy values averaged in 10-bp bins around nucleosome dyads called from chemical cleavage-seq data (Chereji et al. 2018) and averaged over all replicates for the indicated enzymes. qDA-seq data are taken from Chereji et al. (2019b). On the *right*, absolute occupancy values and errors (mean over sites in the bin of the standard deviation among samples) are shown for the maxima and minima of each plot as well as the difference between maximum and minimum values for each enzyme.

and G1-arrested cells (α-factor). However, their qDA-seq AluI data set of exponentially growing cells, that may be biologically closer to our logarithmic phase cells, showed the same trend (Supplemental Fig. S4B). This linker/NFR difference remains to be explained. Nonetheless, the majority of the RE and DNMT data cross-validated each other. Therefore, we were confident that we obtained an accurate measure for absolute occupancy, including the lower linker/NFR occupancy compared to the qDA-seq data.

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

Genome-wide absolute occupancy map and comparison with other nucleosome maps

Because the restriction enzymes were only important for cross-validation but did not contribute significantly to resolution (<1% compared to all DNMT sites) (Fig. 2A; Supplemental Table S1), we used just the ODM-seq absolute occupancy map with a mean resolution of 9 bp and an average error of 6% (mean over sites of standard deviation between samples) (Supplemental Fig. S4C, center and right; Supplemental Table S1) for further analyses. The +1 nucleosome-aligned composite plot of the ODM-seq map (Fig. 1B) was reminiscent of corresponding plots using MNase-seq, MNase-anti-histone-ChIP-seq (Zhang et al. 2011; True et al. 2016), or chemical cleavage-seq data (Fig. 2B), but the other methods show much more pronounced fluctuations in nucleosome peak heights and do not agree, for example, whether the relative occupancies for the +1 nucleosomes are on average higher or lower than those of the +2 nucleosomes (Fig. 2A,B). Our absolute occupancy map now provides meaningful peak heights in absolute terms and thereby resolves such questions; for example, the absolute occupancy for the +1,

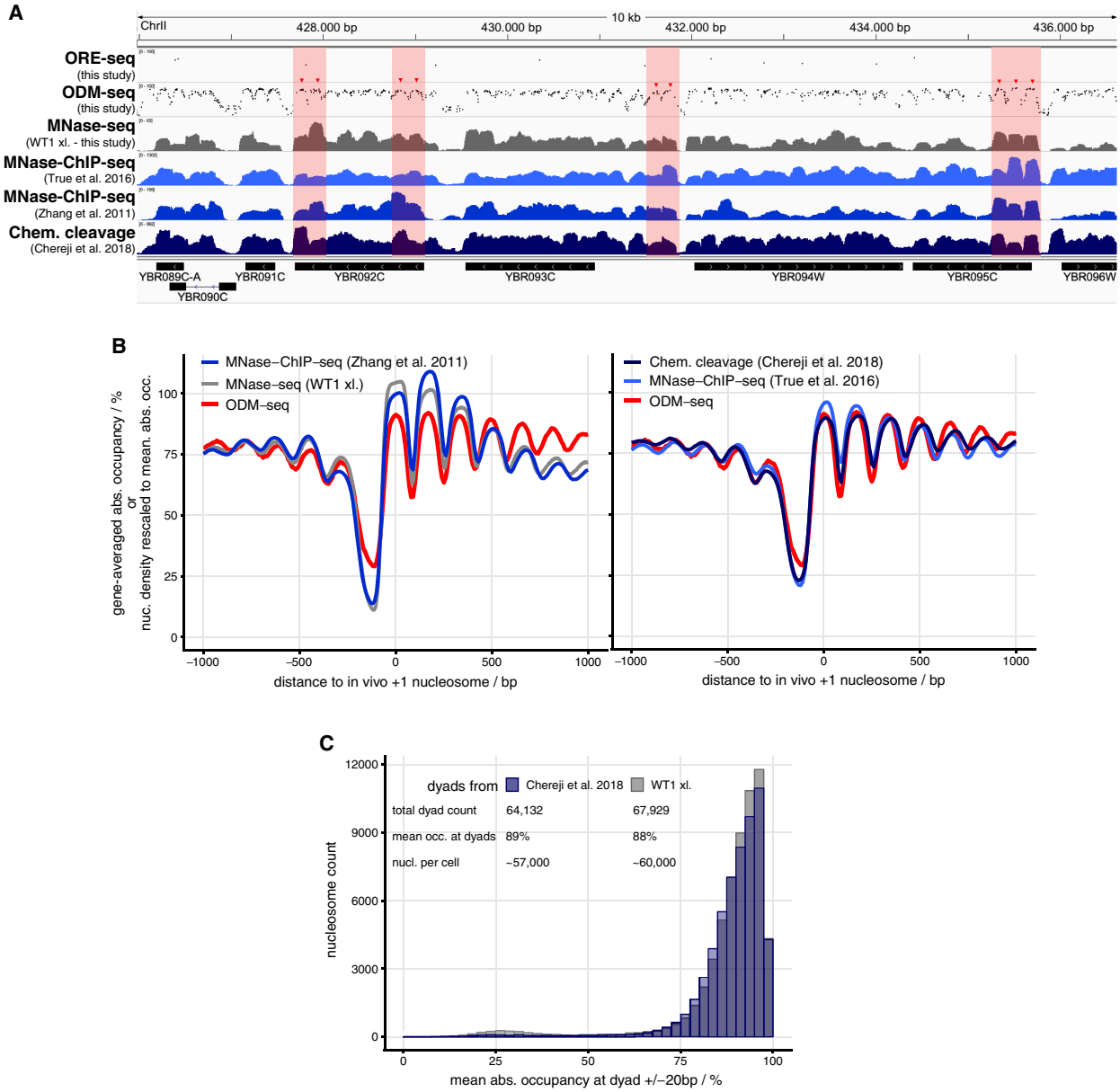


Figure 2. The ODM-seq absolute occupancy map. (*A*) Integrated Genome Viewer (IGV) browser shots comparing the indicated data sets (Supplemental Table S2) with our ORE-seq and ODM-seq absolute occupancy data. Regions in light red highlight pronounced differences in occupancy/signal between methods. (*B*) As in Figure 1C but for the indicated data sets. Because the external data do not provide absolute occupancy, we globally rescaled their signal to have the same genomic mean as the absolute occupancy map. Here and in following cases, nucleosome dyads of external data sets were extended to 147 bp. (*C*) Histogram of absolute occupancy at nucleosome positions called from the indicated data sets.

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

+2, and +3 nucleosomes is on average almost the same. The peak height of nucleosomes further downstream differs slightly if the complete downstream region is plotted or only up to the first transcription termination site (TTS) (Supplemental Fig. S4D) as a promoter NFR follows downstream from a TTS in many cases (Chereji et al. 2017, 2018). Of all compared methods, chemical cleavage-seq was closest to our data in terms of relative peak heights.

We measured generally high occupancy, with a median site occupancy of 84% (Supplemental Fig. S4C, left) but also a substantial fraction of sites with lower occupancies. We underscore that neither ORE-seq nor ODM-seq nor qDA-seq distinguish which kind of factor restricts DNA accessibility, that is, contributes to measured occupancy. To obtain not just an absolute occupancy but more specifically an absolute *nucleosome* occupancy map, our data must be combined with nucleosome-specific mapping data. We called nucleosome dyads either in our own matched MNase-seq data set for the cross-linked WT1 sample or used nucleosome dyad cluster medians (typical nucleosomes) called from chemical cleavage-seq (kindly provided by Razvan Chereji) (Chereji et al. 2018) and determined the absolute occupancy for each of the called dyads or dyad clusters, respectively, by averaging the occupancy of sites within ±20 bp of these calls. The corresponding histograms (Fig. 2C) for chemical cleavage-seq or MNase-seq show rather narrow distributions with means of 89% and 88%, respectively. Dyads with at least 70% absolute occupancy, which correspond to 97% (chemical cleavage) and 94% (MNase-seq) of all dyads that are mappable by ODM-seq have a mean absolute occupancy of 90% (chemical cleavage) and 91% (MNase-seq), both with standard deviation of 6%. The combination of nucleosome calls and absolute occupancy data allows calculating the absolute number of nucleosomes in a yeast cell as ≈57,000 or ≈60,000 nucleosomes per cell for chemical cleavage or MNase-seq calling, respectively. This difference is likely caused by method-specific limitations for scoring nucleosomes. MNase-seq is prone to score nonnucleosomal complexes in NFRs as "nucleosomes" (Chereji et al. 2017), which will lead to an overestimation of nucleosome number and a slight bump in the histogram around 25% occupancy (Fig. 2C). Conversely, high-resolution chemical mapping gives few such false positives but yields clusters that are not well-resolved, especially in gene bodies, that confound the calling algorithm. Therefore, some genic nucleosomes are not called in the latter data set and the number of nucleosomes is underestimated (R Chereji, pers. comm.). This encumbers the exact determination of the number of nucleosomes per cell. Nonetheless, this number likely is in the range of 58,000 ± 1000.

Detection of nonnucleosomal DNA-bound factors by ODM-seq

The more pronounced population of sites with low absolute occupancy around 25% in the histogram of absolute occupancy at CpG/GpC sites (Supplemental Fig. S4C, left) compared to the histogram of absolute occupancies at called nucleosome positions (Fig. 2C) stems mainly from sites in NFRs and linkers (Fig. 1B). NFRs are probably occupied by nonnucleosomal factors (Supplemental Fig. S5A). Indeed, our absolute occupancy map not only shows signals from nucleosomes, but also from DNA binding factors (Gutin et al. 2018) like the general regulatory factors (GRFs) Rap1, Abf1, Mcm1, and Cbf1, but only in few cases for Reb1 and the origin recognition complex subunit Orc1, which may be linked to the mappability of binding site motifs; for example, the Orc1 motif is very AT-rich and hardly mappable by DNMTs (Fig. 3A–D; Supplemental Fig. S5B). Such factor binding is hardly detected, as expected, by MNase-seq or by chemical cleavage (Fig. 3B, center and right). The GRF site-aligned composite plots of absolute occupancy (Fig. 3B, left) show the expected symmetrically aligned and regularly spaced nucleosomal arrays flanking Abf1 and Reb1 sites (Rossi et al. 2018), but much less regular arrays and lower occupancy immediately around Rap1 sites. This reflects that Rap1 sites often come in neighboring pairs, show a broader distribution relative to TSSs (Supplemental Fig. S5A), and occur in unusually wide promoter NFR regions, for example, of the highly expressed ribosomal protein (RP) genes (Knight et al. 2014; Reja et al. 2015).

Correlation of absolute occupancy with transcription rates

Because 97% of all nucleosomes called by chemical mapping have a mean absolute occupancy of 90 ± 6% (±SD) (Fig. 2C), the nucleosome occupancy landscape appeared rather flat and we did not expect strong correlations between absolute occupancies and biological features. Nonetheless, we asked which biological feature correlated with the few percent of nucleosomes that showed lower absolute occupancy.

Very high expression levels, like heat shock–induced genes, were reported to correlate with nucleosome loss over gene bodies (Zhao et al. 2005). We generated composite plots of absolute occupancy, chemical cleavage-seq, and MNase-anti-H3-ChIP-seq (Fig. 4A) as well as other mapping data (Zhang et al. 2011; Joo et al. 2017; Dronamraju et al. 2018; Chereji et al. 2019b) (Supplemental Fig. S6A) for gene quintiles of NET-seq data (nascent RNA bound to RNA polymerase) (Churchman and Weissman 2011), which measures transcription activity. Although MNase-anti-H3-ChIP-seq and ATAC-seq showed reduced signal for the most highly transcribed genes, this was much less the case for chemical cleavage-seq and hardly apparent for the ODM-seq and qDA-seq data. If genes were not binned but if their transcription rates measured by NET-seq or 4sU-seq (Xu et al. 2017) were individually correlated with the various occupancy mapping data (Fig. 4B; Supplemental Fig. S6B), an overall anti-correlation between transcription rates and occupancy was very poor for most and absent for methods that measure absolute occupancy and for chemical cleavage-seq.

Studies from the Pugh (Basehoar et al. 2004), Steinmetz (Xu et al. 2009), and Morillion (van Dijk et al. 2011) groups led to an instructive grouping of transcribed regions in yeast into RP genes, SAGA-, TFIID-dependent genes, cryptic unstable transcripts (CUTs), stable unannotated transcripts (SUTs), and Xrn1-dependent unstable transcripts (XUTs) (Vinayachandran et al. 2018). Our ODM-seq data did not show major differences over gene bodies between these groups (Supplemental Fig. S7A,B), in contrast to relative nucleosome occupancy measurements by MNase-ChIP-seq. Even if these occupancies were rescaled, the distribution of average occupancies across transcripts is too broad for these methods. The high expression levels of RP genes were confirmed by both 4sU-seq and NET-seq, but only 4sU-seq showed much higher transcription rates of SAGA- versus TFIID-dependent genes (Supplemental Fig. S7B).

In addition, we wondered if there was a correlation between the transcribed region length and absolute occupancy. Short units showed the whole range of absolute occupancies, but long units mostly had high absolute occupancy (Supplemental Fig. S6C).

Increased binding of RSC remodeling complex correlates with lower absolute occupancy

Because it was not RNA polymerase passage that reduced absolute occupancy in transcribed regions, we turned to the ATP-

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

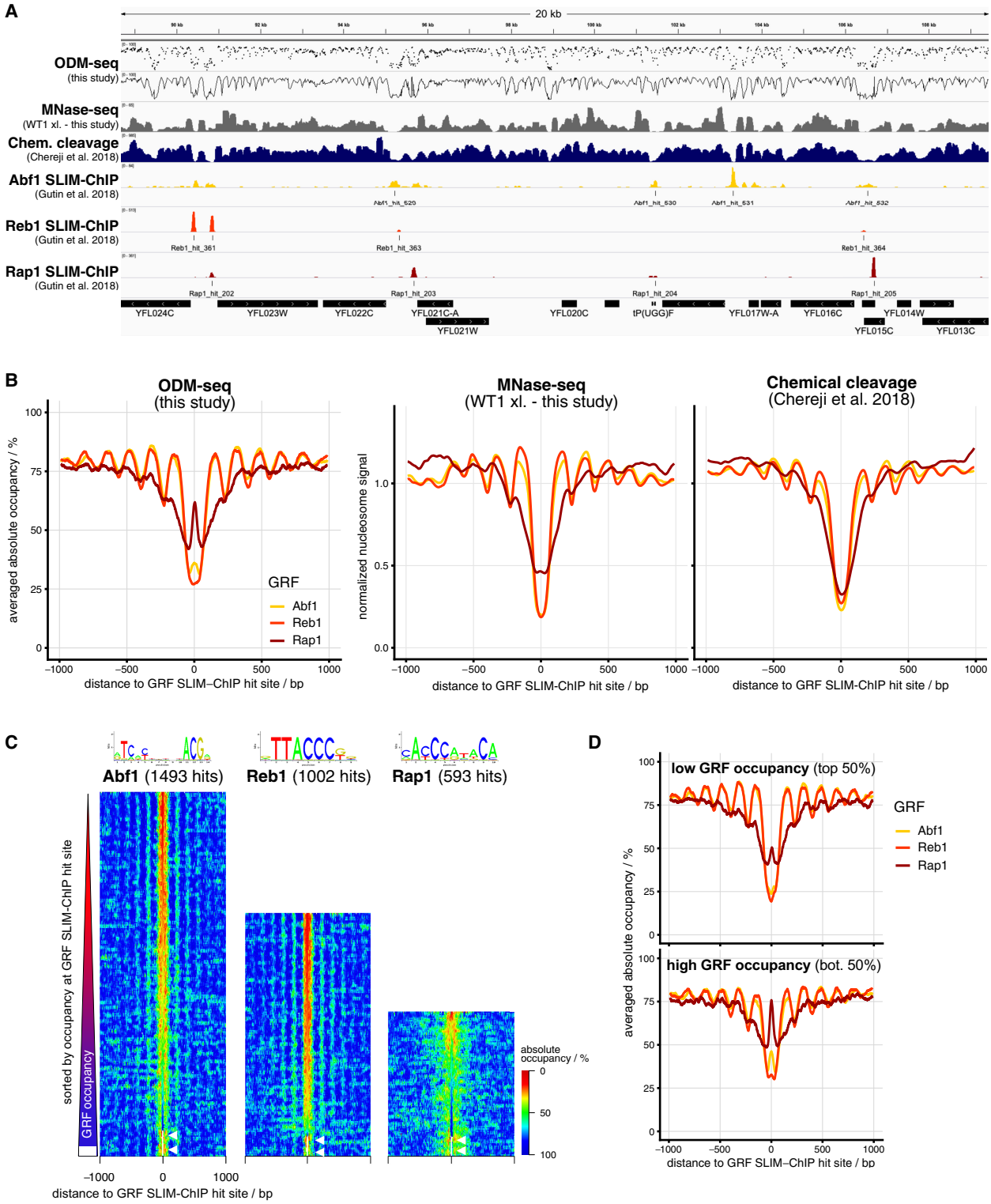


Figure 3. ODM-seq monitors not only absolute nucleosome but also absolute GRF occupancy. (*A*) IGV browser shot comparison of the indicated data sets (Supplemental Table S2). ODM-seq data are given both as individual (*top*) and as connected data points (second from *top*). (*B*) GRF site-aligned composite plots of absolute occupancy (*left*) or normalized signal (*center* and *right*) for the indicated GRFs and data sets. Signals are normalized to a mean of one. (*C*) GRF site-aligned heat maps of absolute occupancy sorted from *top* to *bottom* according to increasing absolute occupancy at GRF sites. The position weight matrix (Badis et al. 2008) and the number of binding sites detected by SLIM-ChIP for the indicated GRFs is given *above* the heat maps. White color denotes absence of signal (highlighted by white arrows). (*D*) As in *B*, *left* graph, but for genes of low and high GRF occupancy according to the SLIM-ChIP sorting in *C*.

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

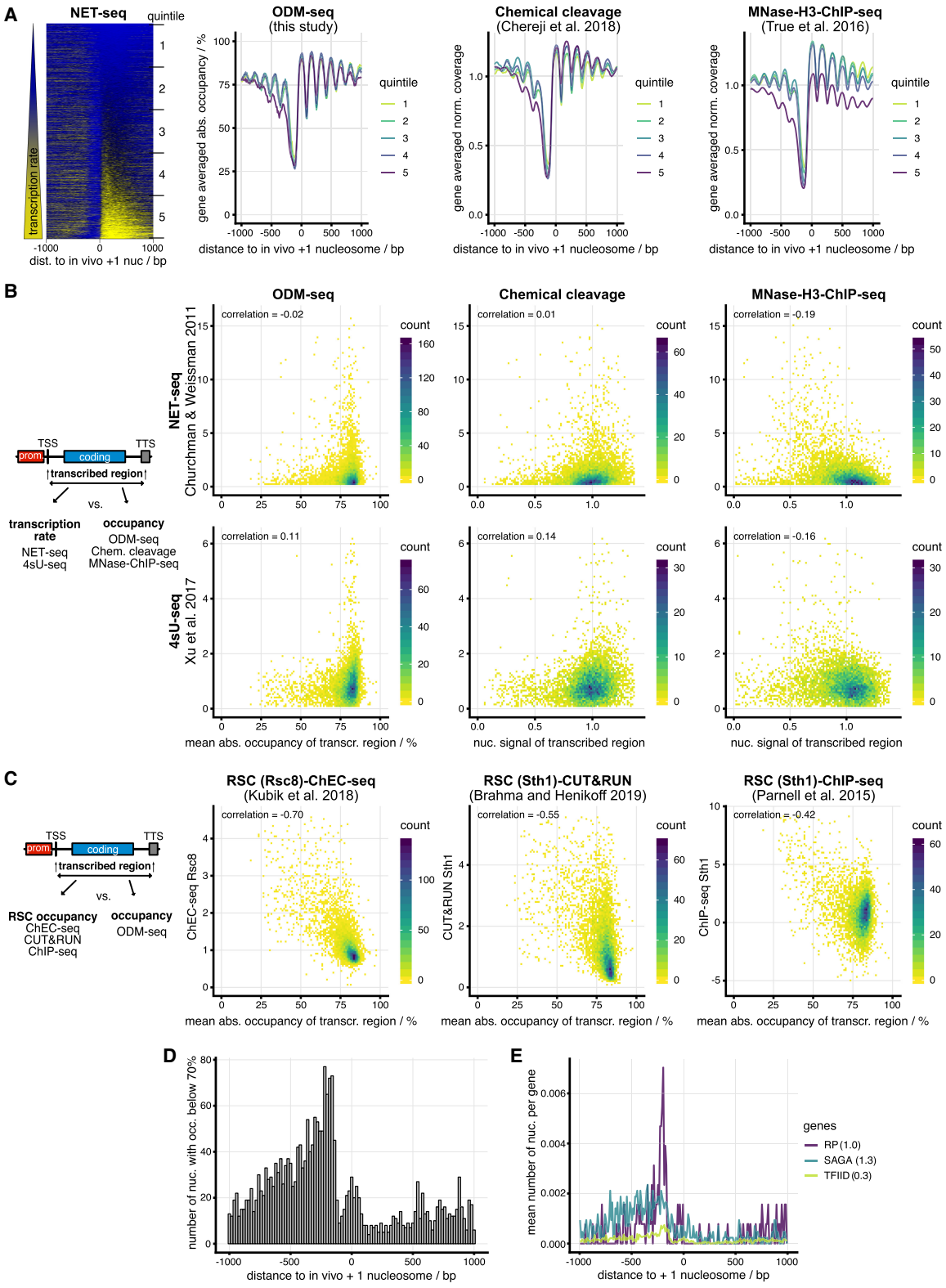


Figure 4. Correlation of absolute occupancy with biological features. (*A, left*) In vivo +1 nucleosome-aligned heat map of NET-seq data monitoring nascent RNA bound to RNA polymerase (Churchman and Weissman 2011) sorted from *top* to *bottom* by increasing signal over the gene body. (*Right*) As in Figure 2B but for the indicated data sets and genes subdivided according to quintiles of sorting in heat map on the *left*. (*B*) Correlation plots (color indicates number of occurrences) of transcription rate (NET-seq as in *A* or 4sU-seq [Xu et al. 2017]) against the absolute occupancy or coverage averaged over transcribed regions for the indicated data sets as in *A*. (*C*) As in *B* but correlation of absolute occupancy averaged over transcribed regions with RSC binding measured by the indicated methods. (*D*) +1 Nucleosome-aligned histogram (accumulated in 20-bp bins) of nucleosomes dyads (Chereji et al. 2018) with <70% absolute occupancy. (*E*) As in *D* but clustered by gene groups (Vinayachandran et al. 2018) as indicated. In brackets, mean number of low absolute occupancy nucleosomes per gene in 2-kb window around in vivo +1 nucleosome. Used data sets are listed in Supplemental Table S2.

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

dependent chromatin remodeling complex RSC, which is the major nucleosome displacing activity in yeast (Cairns et al. 1996; Hartley and Madhani 2009). There was an inverse correlation between RSC signal detected over transcribed regions and their absolute occupancy (Fig. 4C). This was reproducible for different RSC data sets (Parnell et al. 2015; Kubik et al. 2018; Brahma and Henikoff 2019) with different degrees of correlation.

RSC mainly binds and depletes nucleosomes upstream of TSSs (Yen et al. 2012; Krietenstein et al. 2016). Accordingly, those nucleosome dyad clusters—as called by chemical cleavage, which is most reliable in these regions—that had lower than 70% absolute occupancy were mainly upstream of TSSs (Fig. 4D), where regulatory sites, like transcription factor binding sites, are enriched (Lee et al. 2007; Ozonov and van Nimwegen 2013) and mainly for RP and SAGA-dependent genes (Fig. 4E).

The enrichment of low occupancy nucleosomes upstream of TSSs explains in part that the absolute average occupancy upstream of and downstream from NFR minima is 70% and 79%, respectively (Fig. 1B). Nonetheless, this difference is also attributed to genes with wider than average NFRs (less steep upstream flank of the NFR trough in Fig. 1B), so that the alignment at the minimum in the +1 nucleosome-aligned composite plot used for calculating the aforementioned upstream and downstream average occupancies is not in the minimum for these genes. Note that the much lower site average (Fig. 1B, red line) upstream compared to downstream from the NFR is not explained by enrichment of low occupancy nucleosomes, but reflects that upstream nucleosomes are on average less regularly aligned to the +1 nucleosome position than downstream nucleosomes, so that the red line does not mainly average over nucleosome centers but also over linker regions. As shown by the individual data points, upstream nucleosomes also mostly have high absolute occupancy.

We conclude that absolute occupancy is mainly held constant for nucleosomes across the genome and particularly in transcribed regions, unless RSC depletes nucleosomes, which occurs mainly in regulatory regions upstream of TSSs.

Discussion

Here, we present the first genome-wide high-resolution map of absolute occupancy for a eukaryotic genome. Because there was no precedent, we established orthogonal methods, different analysis pipelines, and different experimental conditions that cross-validate each other. During revision of our manuscript, an independent low-resolution approach provided further validation for occupancy at nucleosome dyads (Chereji et al. 2019b). Absolute occupancy was measured as a mirror image of absolute DNA accessibility for REs and DNMTs and cannot distinguish what occupies the DNA. Therefore, we combined our measurements with data from more nucleosome-specific mapping techniques, like MNase-seq or chemical mapping-seq, to arrive at the first high-resolution absolute nucleosome occupancy map.

The main feature of this map is its uniform nucleosome occupancy at dyads. Of all nucleosomes called by chemical mapping, 97% have a mean absolute occupancy >70% with a mean of 90 ± 6% (±SD). This fits to single locus studies; for example, the absolute occupancy of the −2 nucleosome at the repressed yeast *PHO5* promoter was estimated by RE accessibility as 90% (Almer et al. 1986). Nucleosomes are placed along the genome in an all-or-nothing manner without major occupancy differences between the −1, +1, +2, and +3 nucleosomes relative to the TSS. The few nucleosomes with lower absolute occupancy are enriched in regions upstream of TSSs, which fits to high histone turnover (Dion et al. 2007) and abundance of regulatory processes here. High absolute nucleosome occupancy seems to be conserved as qDA-seq in mouse hepatocytes (Chereji et al. 2019b), and quantitative mass spectrometry measurements of histones in *Drosophila* cells (Bonnet et al. 2019) reflected full coverage of the genome with nucleosomes at the species-specific nucleosome repeat length.

The uniformly high absolute nucleosome occupancy suggests that nucleosome depositioning operates as a highly effective default system. This is poorly defined but likely involves histone chaperones, specific histone modifications, like H3K56ac, and remodelers (Almouzni and Cedar 2016). In flies, nucleosomes are rapidly deposited in the wake of DNA replication (Ramachandran and Henikoff 2016) even in regions that are NFRs otherwise. In yeast, promoter NFRs are reestablished almost immediately after replication, maybe by RSC activity, and the kinetics of nucleosome repositioning over genes correlates with transcription levels (Fennessy and Owen-Hughes 2016; Vasseur et al. 2016). Many factors are involved in redepositing nucleosomes in the wake of RNA polymerase (Hennig and Fischer 2013; Smolle et al. 2013). Our finding that highly transcribed genes rarely exhibit low, but mostly exhibit high absolute nucleosome occupancy, argues that RNA polymerase passage, although it requires transient remodeling of nucleosomes (Farnung et al. 2018; Ehara et al. 2019), fosters high nucleosome occupancy, probably via concomitant recruitment of the nucleosome redeposition machinery. It is also compatible with bursty transcription (Haberle and Stark 2018). Conversely, it seems at odds with measurements, for example, by MNase-based methods that suggested an inverse relationship between transcription activity and nucleosome occupancy in gene bodies. MNase is usually not a reliable tool in this regard, as recently reiterated (Chereji et al. 2019a), and methods like DNase-seq or ATAC-seq exaggerate accessibility differences, which was explained recently (Chereji et al. 2019b). Our absolute occupancy measurements clarify this method-driven misconception. Nonetheless, anecdotal reports of nucleosome depletion over highly transcribed genes, like heat shock genes upon heat shock induction (Zhao et al. 2005), were corroborated by qDA-AluI absolute occupancy measurements for the case of a few genes induced by amino acid starvation (Chereji et al. 2019b). So, there may be special cases of highly induced genes that transiently lose nucleosomes over their gene bodies. In addition, regions transcribed by RNA polymerase I and III may be largely nucleosome depleted too, but our analyses focused on genes transcribed by RNA polymerase II.

The minimum of +1 nucleosome-aligned composite absolute occupancy of 29% at promoter NFRs (Fig. 1B) cannot be taken as evidence for higher nucleosome occupancy here than previously thought or for fragile nucleosomes (Kubik et al. 2015; Chereji et al. 2017) but rather reflects binding of nonnucleosomal factors, like transcription factors, GRFs or RNA polymerase. There remains the formal possibility that absolute occupancy, even at nucleosome positions determined by other methods, reflects a composite of nucleosome occupancy and, for example, occupancy by RNA polymerase. However, <1% of yeast genes have more than one RNA polymerase molecule bound (Pelechano et al. 2010), making this formal possibility less of a concern.

Low absolute occupancy correlates with the presence of RSC that is the major nucleosome-ejecting remodeling complex in yeast (Clapier et al. 2016) and particularly responsible for keeping NFRs nucleosome-free (Badis et al. 2008; Hartley and Madhani 2009; Wippo et al. 2011; Krietenstein et al. 2016; Brahma and

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

Henikoff 2019). This correlation immediately suggests a mechanistic explanation for lower nucleosome occupancy, although it remains to be better understood what determines RSC-dependent nucleosome depletion.

Our methodology should be applicable to any chromatin preparation as long as nucleosome dynamics are frozen. We underscore that ORE-seq was necessary for our study as a validating approach but can be omitted in future applications. The application to chromatin of large genomes may be costly because high genome coverage (mean coverage >40-fold) is required and it remains to be tested if cross-linking will be required for stabilizing nucleosomes also in non-yeast species. For sure, care has to be taken to titrate DNA methylation into saturating conditions. The standard NOMe-seq conditions (Kelly et al. 2012) are likely insufficient, especially because DNA methylation by the GpC DNMT is more difficult to saturate.

In summary, DNA methylation under conditions of saturation, frozen nucleosome dynamics, and high sequencing coverage provides a measure for the long-missing chromatin quantity of absolute occupancy for nucleosomes or other factors. This will help distinguish if processes like DNA replication, DNA repair, or aging are associated with changes in nucleosome occupancy.

Methods

Yeast strains and media

The BY4741 strain (MATa *his3Δ0 leu2Δ0 met15Δ0 ura3Δ0*, Euroscarf) was grown to log phase (Supplemental Fig. S1A) in YPDA medium (1% w/v Bacto yeast extract, 2% w/v Bacto peptone, 2% w/v glucose, 0.1 g/L adenine, 1 g/L KH_2PO_4). Cross-linking was with 1% formaldehyde (final concentration) for 1, 5, or 20 min (WT5, other replicates only 20 min) at RT while shaking and quenched for 20 min with 125 (WT1, WT4) or 250 (WT5) mM glycine (final concentration).

Isolation of yeast nuclei

Yeast nuclei were prepared as described (Almer et al. 1986). In brief, cells were harvested by centrifugation, washed in water, resuspended in 2.8 mM EDTA, pH 8, 0.7 M 2-mercaptoethanol, incubated for 30 min at 30°C, washed in 1 M sorbitol, resuspended in 1 M sorbitol, 5 mM 2-mercaptoethanol and spheroplasted by incubation with 20 mg/mL freshly added Zymolyase 100T (MP Biochemicals) for 30 min at 30°C. Spheroplasts were washed with 1 M sorbitol and resuspended in lysis buffer (18% Ficoll, 20 mM KH_2PO_4, 1 mM $MgCl_2$, 0.25 mM EGTA, 0.25 mM EDTA) followed by centrifugation at 22,550*g* for 30 min to collect chromatin ("nuclei"). Pellets were frozen in dry-ice/ethanol and stored at −80°C.

DNA methylation in chromatin

Nuclei pellets were washed in methylation buffer (20 mM HEPES-NaOH pH 7.5, 70 mM NaCl, 0.25 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 0.5% (v/v) glycerol, 1 mM DTT, 0.25 mM PMSF) (Darst et al. 2012). Per reaction, nuclei from approximately 0.1 g wet cell pellet were resuspended in 800 µL methylation buffer containing 640 µM freshly added S-adenosylmethionine (SAM). Two hundred units M.SssI or M.CviPI (both NEB) and 10 mM DTT were freshly added. Methylation reactions were dialyzed in Slide-A-Lyzer MINI Dialysis Devices 10 K MWCO (Thermo Fisher Scientific) against 15 mL methylation buffer with 200 µM freshly added SAM at 25°C for M.SssI or 37°C for M.CviPI. Next, 0.5–1 µg fully assembled pUC19-601-25mer plasmid (pFMP233) (Lowary and Widom 1998; Lieleg et al. 2015a) and *Saccharomyces* Genome Database (SGD) assembled *Escherichia coli* gDNA or SGD assembled *E. coli* plasmid library (limited Sau3A fragments of *E. coli* gDNA ligated into pJET 1.2 plasmid, Thermo Fisher Scientific) was spiked in before methylase addition. *E. coli* spike-in results were not further pursued owing to low coverage. Reactions were stopped by 0.5% SDS (final concentration), reversed cross-linked if required, and DNA was deproteinized by Proteinase K, phenol-chloroform extracted, and RNase A digested.

Restriction enzyme digestion of chromatin

Nuclei from approximately 0.1 g wet cell pellet were prewashed in methylation buffer, centrifuged, and resuspended either in 400 µL 1× CutSmart buffer (NEB) for BamHI-HF and AluI or in 1× NEB2.1 for HindIII. Digestions were started by RE addition, incubated for 30 or 120 min at 37°C, and stopped by 10 mM EDTA and 0.5% SDS (final concentration). DNA preparation was as above. For the cut-all cut method, *Schizosaccharomyces pombe* gDNA, fully digested with the corresponding RE, was spiked in at 5%–10% DNA mass of the final sample before phenol-chloroform extraction. After RNase A digestion, half of the sample was digested for a second time with 100 units of the corresponding RE and in the corresponding buffer for 1.5 h at 37°C. The reaction was stopped by 15 mM EDTA (final concentration). RE accessibility via Southern blot was done as previously described (Musladin et al. 2014).

Calibration samples for restriction enzyme digests

gDNA was purified from BY4741 cells using the Blood & Cell Culture DNA Midi Kit (Qiagen) and one aliquot completely digested with the respective RE. Mixed ratios of digested and not digested gDNA were treated as if they were DNA extracted after restriction enzyme digest of chromatin. The 0% cut calibration sample was an in vivo cross-linked nuclei preparation without RE addition.

DNA methylation and restriction enzyme digestion for in vitro–reconstituted chromatin

SGD chromatin was as in Krietenstein et al. (2012). For low or high nucleosome density, approximately 4 µg or 8 µg *Drosophila* embryo histone octamers, respectively, were assembled with 10 µg yeast plasmid library (Jones et al. 2008). DNA methylation and restriction enzyme digestion were done in the same buffer (20 mM HEPES-NaOH pH 7.5, 70 mM NaCl, 1.5 mM $MgCl_2$, 0.5% [v/v] glycerol, 10 mM DTT) with the same units (20 units or 80 units) for 30 or 180 min at 37°C. For the latter, fresh enzyme (20 units) was refilled after 60 min, in case of DNA methylation, also fresh SAM. Reactions were stopped by 15 mM EDTA and 0.5% SDS (final concentration), followed by Proteinase K digestion and DNA purification. Methylated samples were directly used for library construction. RE digested samples were split after adding *S. pombe* gDNA spike-in, and one-half was digested a second time as described above.

Illumina sequencing library construction and sequencing

Purified DNA was sheared to ~150 bp fragments (Covaris S220) and concentrated (NucleoSpin Gel and PCR Clean-up Kit, Macherey-Nagel). Then, 0.5–1 µg (determined by Qubit, Thermo Fisher Scientific) was used for library preparation: DNA end polishing (15 units T4 DNA Polymerase, 50 units T4 PNK, 5 units Klenow [NEB]) for 30 min at 20°C, DNA purification with AmPure XP beads (Beckman Coulter), A-tailing (15 units Klenow exo$^-$[NEB]) for 30 min at 37°C, AmPure XP bead purification, adapter ligation (15 units T4 DNA Ligase [NEB], 75–150 pmol NEBNext Adapter or

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

NEBNext Methylated Adapter for bisulfite conversion) for 20 min at 25°C, AmPure XP bead purification, PCR (only for half of the sample, 8 PCR cycles with NEBNext Multiplex Primers for Illumina and Phusion Polymerase [NEB]). Methylated samples were either bisulfite converted (Qiagen EpiTect Bisulfite Kit) and subjected to 12–14 PCR cycles (NEBNext Multiplex Primers for Illumina and Phusion U Polymerase, Thermo Fisher Scientific), or treated with the Enzymatic Methyl-Kit (NEB, E7120S). All samples were sequenced on an Illumina HiSeq 1500 in 50-bp paired-end mode.

Oxford Nanopore library construction and sequencing

For Nanopore sequencing, 1 µg of purified DNA was subjected to 1D native barcoding (Oxford Nanopore, SQK-LSK109). To get up to 1 Gb coverage per sample, up to five samples were loaded to a MinION flowcell (R9.4.1).

ORE-seq: cut and uncut fragment count and resection length

After demultiplexing (Girardot et al. 2016), sequenced reads were trimmed based on base-calling quality and mapped (Li and Durbin 2009) using the combined *S. cerevisiae* and *S. pombe* reference genome. Fragments longer than 500 bp or on rDNA loci were removed. At each cut site, fragment starts and ends were counted within a sample-specific window (green areas in Supplemental Fig. S1D; Supplemental Methods) to account for exonuclease resection and within the same window the mean resection length was calculated. Uncut fragments, that is, fragments covering, but not starting or ending at the cut site, were also counted at each cut site. Cut sites with neighbors closer than 200 bp were ignored completely and remaining cut sites with neighbors closer than 300 bp were analyzed only using the starting/ending counts not pointing toward the close neighbor.

ORE-seq: occupancy estimation (cut–uncut method)

Using the cut counts C^i and the uncut counts U^i at cut site i of the sample without a second RE digest, we calculated the effective cut and uncut counts which correct for cut fragment ends caused by shearing and the loss of uncut fragments (detailed derivation in Supplemental Methods):

$$C^i_{\text{eff}} = C^i - \sigma(w+1)U^i\gamma \quad \text{and} \quad U^i_{\text{eff}} = (1+\sigma)U^i\gamma,$$

in which w is the length of the count window, σ is the corrected ratio of all cut counts away from all cut sites, and all uncut fragment counts away from all cut sites and the "uncut correction factor" γ. The estimated occupancy is then given by $U^i_{\text{eff}}/(C^i_{\text{eff}} + U^i_{\text{eff}})$. $C^i_{\text{eff}} + U^i_{\text{eff}}$ denotes the effective coverage and we ignored sites when its value was less than 40. For the uncorrected version, γ is one. In the corrected version, the value of γ was fitted using the calibration samples for AluI, BamHI, and HindIII, minimizing the deviation of the mean occupancy from the prepared occupancy, and varies between 1.555 and 1.680 depending on the enzyme (details in Supplemental Fig. S1G, right, H; Supplemental Methods).

ODM-seq analysis for BS-seq and EM-seq data

We mapped the paired-end reads with BS-Seeker2 (version 2.1.8) (Guo et al. 2013) using Bowtie (Langmead et al. 2009) and trimmed the reads at the real fragment ends by 5 to 10 bp to obtain a constant conversion ratio along reads averaged over all reads, because the conversion ratio usually shows an increase or decrease at these ends owing to end repair. We ignored reads on the loci of rDNA genes and reads with an unconverted HCH motif, because HCH should be fully converted. The "anti-pattern" of the CpG/GpC DNA methyltransferases is the GCH/HCG pattern, respectively, and should also be fully converted. We discarded a sample if the average anti-pattern conversion ratio among all reads was less than 0.98. At each CpG/GpC methylation site, the occupancy is calculated as ratio of the converted reads over the number of all reads. We ignored methylation sites with a coverage less than 20.

ODM-seq analysis for Nanopore-seq data

We called the bases with Guppy (Oxford Nanopore Technologies, version 3.2.2) and mapped the reads with minimap2 (version 2.14-r892-dirty) (Li 2018). For methylation calling we used Nanopolish (version 0.11.0) (Simpson et al. 2017). At each CpG methylation site, the occupancy is calculated as ratio of unmethylated reads over the sum of methylated and unmethylated reads, ignoring reads where the site has been called "ambiguous" by Nanopolish. We ignored methylation sites with a coverage less than 20. Note that, currently, Nanopolish groups CpG sites within 10 bp into one site, thus having a lower resolution than BS-seq.

Calculation of ORE-seq and ODM-seq maps

For RE samples, very rare occupancy estimates outside the interval between 0 and 1 were truncated to 0 or 1. For the ORE-seq map (Fig. 1B), we averaged the occupancy values at the same sites in different RE samples with equal weights. Similarly, for the ODM-seq map (Fig. 1B), different bisulfite samples were averaged with equal weights. We also calculated individual enzyme maps in the same way (Supplemental Table S1).

Bioinformatics

BedGraph files were generated using the R (R Core Team 2018) package rtracklayer (Lawrence et al. 2013) and displayed with Integrative Genomics Viewer (IGV) (Robinson et al. 2011). +1 Nucleosome annotation was generated by calling nucleosomes within a 220-bp window downstream from transcription start sites (Xu et al. 2009) in our MNase-seq WT1 cross-linked data set using DANPOS (Chen et al. 2013). We used the R packages Biostrings (https://rdrr.io/bioc/Biostrings/), GenomicAlignments (Lawrence et al. 2013), and GenomicRanges (Lawrence et al. 2013) to load raw data files.

Data access

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE132225. Source codes are deposited as Supplemental Code and at https://github.com/gerland-group/absolute-occupancy-analysis.

Acknowledgments

We thank Johannes Nübler for pointing us to the question of absolute nucleosome occupancy, Johannes Söding for letting M.H. help start the project, Michael Kladde for advice on DNA methylation of nucleosomes, Tobias Straub and Tamas Schauer for bioinformatics support, Björn Schwalb for sharing 4sU-seq data, Razvan Chereji for sharing chemical cleavage-seq nucleosome cluster calls, Nir Friedman for sharing SLIM-ChIP data, and Daan Verhagen and Felix Müller-Planitz for discussions on DNA methylation footprinting of nucleosomes. This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) via the Collaborative Research Clusters SFB863 and SFB1064. M.W. is a member of the Graduate School of Quantitative

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

Biosciences Munich (QBM). We thank the Amgen Foundation for funding J.L.E.'s stay in the Korber group as an Amgen Scholar.

Author contributions: Conceptualization and methodology, E.O., M.W., N.K., M.H., S.K., H.B., U.G., and P.K.; data curation, formal analysis, and software, E.O., M.W., and M.H.; visualization, E.O. and M.W.; funding acquisition, project administration, supervision, and resources, H.B., U.G. and P.K.; investigation and validation, E.O., N.K., A.S., J.L.E., S.K., and P.K.; writing (original draft), P.K.; writing (review and editing), E.O., M.W., N.K., M.H., U.G., and P.K.

References

Albert I, Mavrich TN, Tomsho LP, Qi J, Zanton SJ, Schuster SC, Pugh BF. 2007. Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* **446:** 572–576. doi:10.1038/nature05632

Almer A, Rudolph H, Hinnen A, Hörz W. 1986. Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements. *EMBO J* **5:** 2689–2696. doi:10.1002/j.1460-2075.1986.tb04552.x

Almouzni G, Cedar H. 2016. Maintenance of epigenetic information. *Cold Spring Harb Perspect Biol* **8:** a019372. doi:10.1101/cshperspect.a019372

Badis G, Chan ET, van Bakel H, Pena-Castillo L, Tillo D, Tsui K, Carlson CD, Gossett AJ, Hasinoff MJ, Warren CL, et al. 2008. A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. *Mol Cell* **32:** 878–887. doi:10.1016/j.molcel.2008.11.020

Barbaric S, Fascher KD, Hörz W. 1992. Activation of the weakly regulated *PHO8* promoter in *S. cerevisiae*: chromatin transition and binding sites for the positive regulatory protein PHO4. *Nucleic Acids Res* **20:** 1031–1038. doi:10.1093/nar/20.5.1031

Bartholomew B. 2014. Regulating the chromatin landscape: structural and mechanistic perspectives. *Annu Rev Biochem* **83:** 671–696. doi:10.1146/annurev-biochem-051810-093157

Basehoar AD, Zanton SJ, Pugh BF. 2004. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **116:** 699–709. doi:10.1016/s0092-8674(04)00205-3.

Bonnet J, Lindeboom RGH, Pokrovsky D, Stricker G, Çelik MH, Rupp RAW, Gagneur J, Vermeulen M, Imhof A, Müller J. 2019. Quantification of proteins and histone marks in *Drosophila* embryos reveals stoichiometric relationships impacting chromatin regulation. *Dev Cell* doi:10.1016/j.devcel.2019.09.011.

Brahma S, Henikoff S. 2019. RSC-associated subnucleosomes define MNase-sensitive promoters in yeast. *Mol Cell* **73:** 238–249.e3. doi:10.1016/j.molcel.2018.10.046

Brogaard K, Xi L, Wang JP, Widom J. 2012. A map of nucleosome positions in yeast at base-pair resolution. *Nature* **486:** 496–501. doi:10.1038/nature11142

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10:** 1213–1218. doi:10.1038/nmeth.2688

Cairns BR, Lorch Y, Li Y, Zhang M, Lacomis L, Erdjument-Bromage H, Tempst P, Du J, Laurent B, Kornberg RD. 1996. RSC, an essential, abundant chromatin-remodeling complex. *Cell* **87:** 1249–1260. doi:10.1016/S0092-8674(00)81820-6

Caserta M, Agricola E, Churcher M, Hiriart E, Verdone L, Di Mauro E, Travers A. 2009. A translational signature for nucleosome positioning *in vivo*. *Nucleic Acids Res* **37:** 5309–5321. doi:10.1093/nar/gkp574

Chen K, Xi Y, Pan X, Li Z, Kaestner K, Tyler J, Dent S, He X, Li W. 2013. DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res* **23:** 341–351. doi:10.1101/gr.142067.112

Chen PB, Zhu LJ, Hainer SJ, McCannell KN, Fazzio TG. 2014. Unbiased chromatin accessibility profiling by RED-seq uncovers unique features of nucleosome variants in vivo. *BMC Genomics* **15:** 1104. doi:10.1186/1471-2164-15-1104

Chereji RV, Ocampo J, Clark DJ. 2017. MNase-sensitive complexes in yeast: nucleosomes and non-histone barriers. *Mol Cell* **65:** 565–577.e3. doi:10.1016/j.molcel.2016.12.009

Chereji RV, Ramachandran S, Bryson TD, Henikoff S. 2018. Precise genome-wide mapping of single nucleosomes and linkers in vivo. *Genome Biol* **19:** 19. doi:10.1186/s13059-018-1398-0

Chereji RV, Bryson TD, Henikoff S. 2019a. Quantitative MNase-seq accurately maps nucleosome occupancy levels. *Genome Biol* **20:** 198. doi:10.1186/s13059-019-1815-z

Chereji RV, Eriksson PR, Ocampo J, Prajapati HK, Clark DJ. 2019b. Accessibility of promoter DNA is not the primary determinant of chromatin-mediated gene regulation. *Genome Res* (this issue). doi:10.1101/gr.249326.119

Churchman LS, Weissman JS. 2011. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469:** 368–373. doi:10.1038/nature09652

Clapier CR, Kasten MM, Parnell TJ, Viswanathan R, Szerlong H, Sirinakis G, Zhang Y, Cairns BR. 2016. Regulation of DNA translocation efficiency within the chromatin remodeler RSC/Sth1 potentiates nucleosome sliding and ejection. *Mol Cell* **62:** 453–461. doi:10.1016/j.molcel.2016.03.032

Cockell M, Rhodes D, Klug A. 1983. Location of the primary sites of micrococcal nuclease cleavage on the nucleosome core. *J Mol Biol* **170:** 423–446. doi:10.1016/S0022-2836(83)80156-9

Cole HA, Howard BH, Clark DJ. 2011. Activation-induced disruption of nucleosome position clusters on the coding regions of Gcn4-dependent genes extends into neighbouring genes. *Nucleic Acids Res* **39:** 9521–9535. doi:10.1093/nar/gkr643

Darst RP, Pardo CE, Ai L, Brown KD, Kladde MP. 2010. Bisulfite sequencing of DNA. *Curr Protoc Mol Biol* **7:** 1–17. doi:10.1002/0471142727.mb0709s91

Darst RP, Nabilsi NH, Pardo CE, Riva A, Kladde MP. 2012. DNA methyltransferase accessibility protocol for individual templates by deep sequencing. *Meth Enzymol* **513:** 185–204. doi:10.1016/B978-0-12-391938-0.00008-2

DeGennaro CM, Alver BH, Marguerat S, Stepanova E, Davis CP, Bahler J, Park PJ, Winston F. 2013. Spt6 regulates intragenic and antisense transcription, nucleosome positioning, and histone modifications genome-wide in fission yeast. *Mol Cell Biol* **33:** 4779–4792. doi:10.1128/MCB.01068-13

Dingwall C, Lomonossoff GP, Laskey RA. 1981. High sequence specificity of micrococcal nuclease. *Nucleic Acids Res* **9:** 2659–2673. doi:10.1093/nar/9.12.2659

Dion MF, Kaplan T, Kim M, Buratowski S, Friedman N, Rando OJ. 2007. Dynamics of replication-independent histone turnover in budding yeast. *Science* **315:** 1405–1408. doi:10.1126/science.1134053

Dronamraju R, Hepperla AJ, Shibata Y, Adams AT, Magnuson T, Davis IJ, Strahl BD. 2018. Spt6 association with RNA polymerase II directs mRNA turnover during transcription. *Mol Cell* **70:** 1054–1066. doi:10.1016/j.molcel.2018.05.020

Ehara H, Kujirai T, Fujino Y, Shirouzu M, Kurumizaka H, Sekine SI. 2019. Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science* **363:** 744–747. doi:10.1126/science.aav8912

Farnung L, Vos SM, Cramer P. 2018. Structure of transcribing RNA polymerase II-nucleosome complex. *Nat Commun* **9:** 5432. doi:10.1038/s41467-018-07870-y

Fennessy RT, Owen-Hughes T. 2016. Establishment of a promoter-based chromatin architecture on recently replicated DNA can accommodate variable inter-nucleosome spacing. *Nucleic Acids Res* **44:** 7189–7203. doi:10.1093/nar/gkw331

Flaus A, Luger K, Tan S, Richmond TJ. 1996. Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. *Proc Natl Acad Sci* **93:** 1370–1375. doi:10.1073/pnas.93.4.1370

Flores O, Deniz O, Soler-López M, Orozco M. 2014. Fuzziness and noise in nucleosomal architecture. *Nucleic Acids Res* **42:** 4934–4946. doi:10.1093/nar/gku165

Gargiulo G, Levy S, Bucci G, Romanenghi M, Fornasari L, Beeson KY, Goldberg SM, Cesaroni M, Ballarini M, Santoro F, et al. 2009. NA-Seq: a discovery tool for the analysis of chromatin structure and dynamics during differentiation. *Dev Cell* **16:** 466–481. doi:10.1016/j.devcel.2009.02.002

Girardot C, Scholtalbers J, Sauer S, Su SY, Furlong EE. 2016. Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics* **17:** 419. doi:10.1186/s12859-016-1284-2

Givens RM, Lai WK, Rizzo JM, Bard JE, Mieczkowski PA, Leatherwood J, Huberman JA, Buck MJ. 2012. Chromatin architectures at fission yeast transcriptional promoters and replication origins. *Nucleic Acids Res* **40:** 7176–7189. doi:10.1093/nar/gks351

Gregory PD, Barbaric S, Horz W. 1999. Restriction nucleases as probes for chromatin structure. *Methods Mol Biol* **119:** 417–425. doi:10.1385/1-59259-681-9:417

Guo W, Fiziev P, Yan W, Cokus S, Sun X, Zhang MQ, Chen PY, Pellegrini M. 2013. BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. *BMC Genomics* **14:** 774. doi:10.1186/1471-2164-14-774

Gutin J, Sadeh R, Bodenheimer N, Joseph-Strauss D, Klein-Brill A, Alajem A, Ram O, Friedman N. 2018. Fine-resolution mapping of TF binding and chromatin interactions. *Cell Rep* **22:** 2797–2807. doi:10.1016/j.celrep.2018.02.052

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

Haberle V, Stark A. 2018. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol* **19:** 621–637. doi:10.1038/s41580-018-0028-8

Hartley PD, Madhani HD. 2009. Mechanisms that specify promoter nucleosome location and identity. *Cell* **137:** 445–458. doi:10.1016/j.cell.2009.02.043

Hennig BP, Fischer T. 2013. The great repression: chromatin and cryptic transcription. *Transcription* **4:** 97–101. doi:10.4161/trns.24884

Hörz W, Altenburger W. 1981. Sequence specific cleavage of DNA by micrococcal nuclease. *Nucleic Acids Res* **9:** 2643–2658. doi:10.1093/nar/9.12.2643

Jessen WJ, Dhasarathy A, Hoose SA, Carvin CD, Risinger AL, Kladde MP. 2004. Mapping chromatin structure in vivo using DNA methyltransferases. *Methods* **33:** 68–80. doi:10.1016/j.ymeth.2003.10.025

Jiang C, Pugh BF. 2009. Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* **10:** 161–172. doi:10.1038/nrg2522

Jones GM, Stalker J, Humphray S, West A, Cox T, Rogers J, Dunham I, Prelich G. 2008. A systematic library for comprehensive overexpression screens in *Saccharomyces cerevisiae*. *Nat Methods* **5:** 239–241. doi:10.1038/nmeth.1181

Joo YJ, Ficarro SB, Soares LM, Chun Y, Marto JA, Buratowski S. 2017. Downstream promoter interactions of TFIID TAFs facilitate transcription reinitiation. *Genes Dev* **31:** 2162–2174. doi:10.1101/gad.306324.117

Kaplan N, Hughes TR, Lieb JD, Widom J, Segal E. 2010. Contribution of histone sequence preferences to nucleosome organization: proposed definitions and methodology. *Genome Biol* **11:** 140. doi:10.1186/gb-2010-11-11-140

Kelly TK, Liu Y, Lay FD, Liang G, Berman BP, Jones PA. 2012. Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res* **22:** 2497–2506. doi:10.1101/gr.143008.112

Kilgore JA, Hoose SA, Gustafson TL, Porter W, Kladde MP. 2007. Single-molecule and population probing of chromatin structure using DNA methyltransferases. *Methods* **41:** 320–332. doi:10.1016/j.ymeth.2006.08.008

Knight B, Kubik S, Ghosh B, Bruzzone MJ, Geertz M, Martin V, Dénervaud N, Jacquet P, Ozkan B, Rougemont J, et al. 2014. Two distinct promoter architectures centered on dynamic nucleosomes control ribosomal protein gene transcription. *Genes Dev* **28:** 1695–1709. doi:10.1101/gad.244434.114

Korolev N, Vorontsova OV, Nordenskiöld L. 2007. Physicochemical analysis of electrostatic foundation for DNA–protein interactions in chromatin transformations. *Prog Biophys Mol Biol* **95:** 23–49. doi:10.1016/j.pbiomolbio.2006.11.003

Krebs AR, Imanci D, Hoerner L, Gaidatzis D, Burger L, Schübeler D. 2017. Genome-wide single-molecule footprinting reveals high RNA polymerase II turnover at paused promoters. *Mol Cell* **67:** 411–422.e4. doi:10.1016/j.molcel.2017.06.027

Krietenstein N, Wippo CJ, Lieleg C, Korber P. 2012. Genome-wide *in vitro* reconstitution of yeast chromatin with *in vivo*-like nucleosome positioning. *Meth Enzymol* **513:** 205–232. doi:10.1016/B978-0-12-391938-0.00009-4

Krietenstein N, Wal M, Watanabe S, Park B, Peterson CL, Pugh BF, Korber P. 2016. Genomic nucleosome organization reconstituted with pure proteins. *Cell* **167:** 709–721.e12. doi:10.1016/j.cell.2016.09.045

Kubik S, Bruzzone MJ, Jacquet P, Falcone JL, Rougemont J, Shore D. 2015. Nucleosome stability distinguishes two different promoter types at all protein-coding genes in yeast. *Mol Cell* **60:** 422–434. doi:10.1016/j.molcel.2015.10.002

Kubik S, O'Duibhir E, de Jonge WJ, Mattarocci S, Albert B, Falcone JL, Bruzzone MJ, Holstege FCP, Shore D. 2018. Sequence-directed action of RSC remodeler and general regulatory factors modulates +1 nucleosome position to facilitate transcription. *Mol Cell* **71:** 89–102. doi:10.1016/j.molcel.2018.05.030

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25. doi:10.1186/gb-2009-10-3-r25

Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. 2013. Software for computing and annotating genomic ranges. *PLoS Comput Biol* **9:** e1003118. doi:10.1371/journal.pcbi.1003118

Lee W, Tillo D, Bray N, Morse RH, Davis RW, Hughes TR, Nislow C. 2007. A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* **39:** 1235–1244. doi:10.1038/ng2117

Levo M, Avnit-Sagi T, Lotan-Pompan M, Kalma Y, Weinberger A, Yakhini Z, Segal E. 2017. Systematic investigation of transcription factor activity in the context of chromatin using massively parallel binding and expression assays. *Mol Cell* **65:** 604–617.e6. doi:10.1016/j.molcel.2017.01.007

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34:** 3094–3100. doi:10.1093/bioinformatics/bty191

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25:** 1754–1760. doi:10.1093/bioinformatics/btp324

Lieleg C, Ketterer P, Nuebler J, Ludwigsen J, Gerland U, Dietz H, Mueller-Planitz F, Korber P. 2015a. Nucleosome spacing generated by ISWI and CHD1 remodelers is constant regardless of nucleosome density. *Mol Cell Biol* **35:** 1588–1605. doi:10.1128/MCB.01070-14

Lieleg C, Krietenstein N, Walker M, Korber P. 2015b. Nucleosome positioning in yeasts: methods, maps, and mechanisms. *Chromosoma* **124:** 131–151. doi:10.1007/s00412-014-0501-x

Lowary PT, Widom J. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* **276:** 19–42. doi:10.1006/jmbi.1997.1494

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389:** 251–260. doi:10.1038/38444

Meyer CA, Liu XS. 2014. Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. *Nat Rev Genet* **15:** 709–721. doi:10.1038/nrg3788

Moyle-Heyrman G, Zaichuk T, Xi L, Zhang Q, Uhlenbeck OC, Holmgren R, Widom J, Wang JP. 2013. Chemical map of *Schizosaccharomyces pombe* reveals species-specific features in nucleosome positioning. *Proc Natl Acad Sci* **110:** 20158–20163. doi:10.1073/pnas.1315809110

Musladin S, Krietenstein N, Korber P, Barbaric S. 2014. The RSC chromatin remodeling complex has a crucial role in the complete remodeler set for yeast *PHO5* promoter opening. *Nucleic Acids Res* **42:** 4270–4282. doi:10.1093/nar/gkt1395

Noll M. 1974. Subunit structure of chromatin. *Nature* **251:** 249–251. doi:10.1038/251249a0

Ozonov EA, van Nimwegen E. 2013. Nucleosome free regions in yeast promoters result from competitive binding of transcription factors that interact with chromatin modifiers. *PLoS Comput Biol* **9:** e1003181. doi:10.1371/journal.pcbi.1003181

Parnell TJ, Schlichter A, Wilson BG, Cairns BR. 2015. The chromatin remodelers RSC and ISW1 display functional and chromatin-based promoter antagonism. *eLife* **4:** e06073. doi:10.7554/eLife.06073

Pelechano V, Chávez S, Pérez-Ortín JE. 2010. A complete set of nascent transcription rates for yeast genes. *PLoS One* **5:** e15442. doi:10.1371/journal.pone.0015442

Polach KJ, Widom J. 1995. Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J Mol Biol* **254:** 130–149. doi:10.1006/jmbi.1995.0606

Quintales L, Vázquez E, Antequera F. 2015. Comparative analysis of methods for genome-wide nucleosome cartography. *Brief Bioinform* **16:** 576–587. doi:10.1093/bib/bbu037

R Core Team. 2018. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/.

Ramachandran S, Henikoff S. 2016. Transcriptional regulators compete with nucleosomes post-replication. *Cell* **165:** 580–592. doi:10.1016/j.cell.2016.02.062

Reja R, Vinayachandran V, Ghosh S, Pugh BF. 2015. Molecular mechanisms of ribosomal protein gene coregulation. *Genes Dev* **29:** 1942–1954. doi:10.1101/gad.268896.115

Rill R, Van Holde KE. 1973. Properties of nuclease-resistant fragments of calf thymus chromatin. *J Biol Chem* **248:** 1080–1083.

Rizzo JM, Mieczkowski PA, Buck MJ. 2011. Tup1 stabilizes promoter nucleosome positioning and occupancy at transcriptionally plastic genes. *Nucleic Acids Res* **39:** 8803–8819. doi:10.1093/nar/gkr557

Rizzo JM, Bard JE, Buck MJ. 2012. Standardized collection of MNase-seq experiments enables unbiased data set comparisons. *BMC Mol Biol* **13:** 15. doi:10.1186/1471-2199-13-15

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29:** 24–26. doi:10.1038/nbt.1754

Rossi MJ, Lai WKM, Pugh BF. 2018. Genome-wide determinants of sequence-specific DNA binding of general regulatory factors. *Genome Res* **28:** 497–508. doi:10.1101/gr.229518.117

Seeber A, Hauer MH, Gasser SM. 2018. Chromosome dynamics in response to DNA damage. *Annu Rev Genet* **52:** 295–319. doi:10.1146/annurev-genet-120417-031334

Simpson JT, Workman RE, Zuzarte PC, David M, Dursi LJ, Timp W. 2017. Detecting DNA cytosine methylation using nanopore sequencing. *Nat Methods* **14:** 407–410. doi:10.1038/nmeth.4184

Small EC, Xi L, Wang JP, Widom J, Licht JD. 2014. Single-cell nucleosome mapping reveals the molecular basis of gene expression heterogeneity. *Proc Natl Acad Sci* **111:** E2462–E2471. doi:10.1073/pnas.1400517111

Smolle M, Workman JL, Venkatesh S. 2013. reSETting chromatin during transcription elongation. *Epigenetics* **8:** 10–15. doi:10.4161/epi.23333

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

Stein A, Takasuka TE, Collings CK. 2010. Are nucleosome positions *in vivo* primarily determined by histone–DNA sequence preferences? *Nucleic Acids Res* **38:** 709–719. doi:10.1093/nar/gkp1043

True JD, Muldoon JJ, Carver MN, Poorey K, Shetty SJ, Bekiranov S, Auble DT. 2016. The modifier of transcription 1 (Mot1) ATPase and Spt16 histone chaperone co-regulate transcription through preinitiation complex assembly and nucleosome organization. *J Biol Chem* **291:** 15307–15319. doi:10.1074/jbc.M116.735134

van Dijk EL, Chen CL, d'Aubenton-Carafa Y, Gourvennec S, Kwapisz M, Roche V, Bertrand C, Silvain M, Legoix-Né P, Loeillet S, et al. 2011. XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* **475:** 114–117. doi:10.1038/nature10118

Vasseur P, Tonazzini S, Ziane R, Camasses A, Rando OJ, Radman-Livaja M. 2016. Dynamics of nucleosome positioning maturation following genomic replication. *Cell Rep* **16:** 2651–2665. doi:10.1016/j.celrep.2016.07.083

Venkatesh S, Workman JL. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol* **16:** 178–189. doi:10.1038/nrm3941

Vinayachandran V, Reja R, Rossi MJ, Park B, Rieber L, Mittal C, Mahony S, Pugh BF. 2018. Widespread and precise reprogramming of yeast protein-genome interactions in response to heat shock. *Genome Res* **28:** 357–366. doi:10.1101/gr.226761.117.

Wal M, Pugh BF. 2012. Genome-wide mapping of nucleosome positions in yeast using high-resolution MNase ChIP-Seq. *Meth Enzymol* **513:** 233–250. doi:10.1016/B978-0-12-391938-0.00010-0

Weiner A, Hughes A, Yassour M, Rando OJ, Friedman N. 2010. High-resolution nucleosome mapping reveals transcription-dependent promoter packaging. *Genome Res* **20:** 90–100. doi:10.1101/gr.098509.109

Wippo CJ, Israel L, Watanabe S, Hochheimer A, Peterson CL, Korber P. 2011. The RSC chromatin remodelling enzyme has a unique role in directing the accurate positioning of nucleosomes. *EMBO J* **30:** 1277–1288. doi:10.1038/emboj.2011.43

Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM. 2009. Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457:** 1033–1037. doi:10.1038/nature07728

Xu Y, Bernecky C, Lee CT, Maier KC, Schwalb B, Tegunov D, Plitzko JM, Urlaub H, Cramer P. 2017. Architecture of the RNA polymerase II-Paf1C-TFIIS transcription elongation complex. *Nat Commun* **8:** 15741. doi:10.1038/ncomms15741

Yen K, Vinayachandran V, Batta K, Koerber RT, Pugh BF. 2012. Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell* **149:** 1461–1473. doi:10.1016/j.cell.2012.04.036

Yuan GC, Liu YJ, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ. 2005. Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* **309:** 626–630. doi:10.1126/science.1112178

Zhang Y, Moqtaderi Z, Rattner BP, Euskirchen G, Snyder M, Kadonaga JT, Liu XS, Struhl K. 2009. Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions *in vivo*. *Nat Struct Mol Biol* **16:** 847–852. doi:10.1038/nsmb.1636

Zhang Z, Wippo CJ, Wal M, Ward E, Korber P, Pugh BF. 2011. A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* **332:** 977–980. doi:10.1126/science.1200508

Zhao J, Herrera-Diaz J, Gross DS. 2005. Domain-wide displacement of histones by activated heat shock factor occurs independently of Swi/Snf and is not correlated with RNA polymerase II density. *Mol Cell Biol* **25:** 8985–8999. doi:10.1128/MCB.25.20.8985-8999.2005

Zhou CY, Johnson SL, Gamarra NI, Narlikar GJ. 2016. Mechanisms of ATP-dependent chromatin remodeling motors. *Annu Rev Biophys* **45:** 153–181. doi:10.1146/annurev-biophys-051013-022819

Received June 7, 2019; accepted in revised form October 31, 2019.

Downloaded from genome.cshlp.org on November 9, 2020 - Published by Cold Spring Harbor Laboratory Press

Absolute nucleosome occupancy map for the *Saccharomyces cerevisiae* genome

Elisa Oberbeckmann, Michael Wolff, Nils Krietenstein, et al.

Genome Res. 2019 29: 1996-2009 originally published online November 6, 2019
Access the most recent version at doi:10.1101/gr.253419.119

Supplemental Material http://genome.cshlp.org/content/suppl/2019/11/25/gr.253419.119.DC1

Related Content **Accessibility of promoter DNA is not the primary determinant of chromatin-mediated gene regulation**
Razvan V. Chereji, Peter R. Eriksson, Josefina Ocampo, et al.
Genome Res. December , 2019 29: 1985-1995

References This article cites 105 articles, 22 of which can be accessed free at:
http://genome.cshlp.org/content/29/12/1996.full.html#ref-list-1

Articles cited in:
http://genome.cshlp.org/content/29/12/1996.full.html#related-urls

Creative Commons License This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

To subscribe to *Genome Research* go to:
http://genome.cshlp.org/subscriptions

© 2019 Oberbeckmann et al.; Published by Cold Spring Harbor Laboratory Press

CHAPTER 2: THE NUCLEAR ACTIN-CONTAINING ARP8 MODULE IS A LINKER DNA SENSOR DRIVING INO80 CHROMATIN REMODELING

Kilian R. Knoll[1,2,9], Sebastian Eustermann[1,2,9], Vanessa Niebauer[1,2], Elisa Oberbeckmann[3], Gabriele Stoehr[1,2,7], Kevin Schall[1,2], Alessandro Tosi[1,2,8], Marianne Schwarz[1,2,4], Andrea Buchfellner[5], Philipp Korber[3] and Karl-Peter Hopfner[1,2,6]

[1]Department of Biochemistry, Ludwig-Maximilians-Universität München, Munich, Germany. [2]Gene Center, Ludwig-Maximilians-Universität München, Munich, Germany. [3]Chair of Molecular Biology, Biomedical Center, Faculty of Medicine, Ludwig-Maximilians-Universität München, Munich, Germany. [4]Institute of Biophysics, Ulm University, Ulm, Germany. [5]ChromoTek GmbH, Planegg, Germany. [6]Center for Integrated Protein Science, Ludwig-Maximilians-Universität München, Munich, Germany. [7]Present address: OmicScouts GmbH, Freising, Germany. [8]Present address: Vossius & Partner, Munich, Germany. [9]These authors contributed equally to the work.

This chapter was published in Nature Structure and Molecular Biology. Due to publishing permissions, the accepted manuscript was printed here.
DOI: 10.1038/s41594-018-0115-8

Author contributions to "The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling"

K.R.K. and S.E. determined the structures and built atomic models; K.R.K. prepared samples for crystallization and performed biochemical analysis of the Arp8 module; S.E., A.T., M.S., and A.B identified the Arp4–N-actin binding nanobody and performed its initial characterization; K.R.K. and G.S. performed affinity enrichment mass spectrometry analysis; S.E. and K.-P.H. devised with a contribution of M.S. preparation and characterization of recombinant INO80 complex; V.N. prepared mutant complexes and performed their biochemical analysis; V.N. and me performed and analyzed genome-wide remodeling assays under supervision of P.K.; K.S. prepared nucleosomes; S.E. and K.-P.H. designed the overall study, analyzed the results, and wrote the paper with contributions from K.R.K., V.N., P.K. and me;

The nuclear actin-Arp4-Arp8-module is a linker DNA sensor driving INO80 chromatin remodelling

Author list and affiliations

Kilian R. Knoll[1,2,9], Sebastian Eustermann[1,2,9], Vanessa Niebauer[1,2], Elisa Oberbeckmann[3], Gabriele Stoehr[1,2,7], Kevin Schall[1,2], Alessandro Tosi[1,2,8], Marianne Schwarz[1,2,4], Andrea Buchfellner[5], Philipp Korber[3] & Karl-Peter Hopfner[1,2,6,10].

[1]Department of Biochemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany.
[2]Gene Center, Ludwig-Maximilians-Universität München, 81377 Munich, Germany.
[3]Chair of Molecular Biology, Biomedical Center (BMC), Faculty of Medicine, LMU Munich, 82152 Martinsried near Munich, Germany.
[4]Institute of Biophysics, Ulm University, 89081 Ulm, Germany.
[5]ChromoTek GmbH, 82152 Planegg, Germany.
[6]Center for Integrated Protein Science (CIPSM), Ludwig-Maximilians-Universität München, 81377 Munich, Germany.
[7]Present address: OmicScouts GmbH, 85354 Freising, Germany.
[8]Present address: VOSSIUS & PARTNER, 81675 Munich, Germany.
[9]These authors contributed equally to this work.
[10]Correspondence should be addressed to K.P.H (hopfner@genzentrum.lmu.de).

Abstract

Nuclear actin (N-actin) and actin-related proteins (Arps) are critical components of several chromatin modulating complexes, including the chromatin remodeller INO80, but their function is largely elusive. Here, we report the crystal structure of the 180 kDa Arp8-module of *S. cerevisiae* INO80 complex and establish its role in recognition of extranucleosomal linker DNA. Arp8 engages N-actin unlike other known interactions between actin-fold proteins and specifies thereby recruitment of the Arp4-N-actin heterodimer to a segmented scaffold of the helicase-SANT–associated (HSA) domain of Ino80. The helical HSA domain spans a distance over 120 Å and provides, adjacent to the Snf2-type ATPase motor, an extended binding platform for extranucleosomal DNA, which is required for nucleosome sliding and genome-wide nucleosome positioning. Together with the recent cryoEM structure of INO80^{Core}-nucleosome complex, our findings suggest an allosteric mechanism, by which INO80 senses 40 bp extranucleosomal entry DNA to conduct highly processive chromatin remodelling.

Introduction

ATP-dependent chromatin remodellers shape the spatial and temporal organization of chromatin and generate hallmark features such as regularly spaced nucleosomal arrays flanking nucleosome depleted regions (NDRs) at promoters[1,2]. Remodellers are generally grouped into four families INO80, SWI/SNF, ISWI and CHD, according to sequence similarities in their common Snf2-type ATPase motor domain. They utilize ATP-dependent DNA translocation to catalyse different types of large-scale nucleosome remodelling reactions – sliding, eviction/assembly, positioning and editing (histone exchange)[3,4].

INO80 and SWI/SNF family remodellers are mega-Dalton complexes comprising typically more than 15 different protein subunits[4]. A unifying but poorly understood key feature of these two multi-subunit remodeller families is the presence of nuclear actin (N-actin) and actin related proteins (Arps). *S. cerevisiae* possesses altogether ten Arps. Arp4-9 localize to the nucleus as integral, functionally important subunits of INO80 and SWI/SNF remodellers and of the histone acetyl transferase NuA4/TIP60[4-8]. Arp4 and N-actin form an evolutionarily conserved pair in all of these enzymes, except yeast SWI/SNF and RSC, where the Arp4-N-actin-pair is replaced by the diverged, but structurally related Arp7-Arp9 pair. Structural studies of Arp4-N-actin or Arp7-Arp9 revealed binding via their barbed ends to a helical HSA domain N-terminal to the Snf2-type ATPase domain of Swr1 and Sth1, respectively[9-11]. N-actin and nuclear Arps play an essential role in cellular stress response as well as during development[12,13], and respective genes, encoding, for example, the human Arp4 homolog BAF53, are frequently mutated in cancer[13,14]. However, the precise molecular mechanism explaining the functional importance of N-actin and nuclear Arps remains still largely elusive.

The INO80 complex is particularly intriguing for studying the functional role of nuclear actin-fold proteins[6,15]. In addition to N-actin and Arp4, INO80 contains with Arp5 and Arp8 in total four actin-fold proteins and is conserved in this respect from yeast to man [6,16]. INO80 has pivotal functions in gene regulation, replication and genome maintenance[16,17], as it slides[15], edits[18] and positions[2,19] nucleosomes including the +1 nucleosome at promoter regions[2]. INO80 has a modular architecture[11,16,20,21]. The Ino80 protein subunit, harbouring the Snf2-type ATPase motor, is an assembly platform for the other subunits: its N-terminal region interacts in yeast with the species-specific “Nhp10-module” (a subcomplex of Ino eighty subunits Ies1, Ies3, Ies5, and Nhp10), which regulates the switch-like stimulation of INO80’s nucleosome sliding efficiency by extranucleosomal DNA > 40 base pairs[22]. The middle region of Ino80 contains the HSA domain (Ino80^{HSA}), which binds the highly conserved “Arp8-module” composed of N-actin, Arp4, Arp8, Ies4 and Taf14[11,16]. Deletion of Arp8 or the HSA-domain leads to the loss of the whole Arp8-module and results in a remodelling defective

INO80 complex[15,16,23]. The C-terminal region of Ino80 forms the equally conserved INO80 core module (INO80^{Core}), containing the Snf2-type ATPase, Ies2, the Arp5-Ies6 complex and the Rvb1-Rvb2 heterohexameric AAA-type ATPases. The structure and unified mechanism by which INO80^{Core} recognises and remodels the nucleosome core particle (NCP) has been recently revealed at high resolution by cryo-EM[24,25]. We uncovered also that the function of INO80^{Core} as a macromolecular ratchet depends critically on a direct interaction of Arp5 with nucleosomal DNA[24]. It has been proposed that other nuclear Arps could be involved in DNA or nucleosome interactions[23,26,27]. Indeed, our cryoEM analysis of a fungal INO80 complex, which included all evolutionarily conserved subunits, located the Arp8-module near extranucleosomal entry DNA, but its analysis was, unlike to the NCP-INO80^{Core} region of the complex, limited by lower resolution. Until now, high-resolution structural information on the functionally critical architecture of the Arp8-module is missing.

Here, we report the crystal structure of the Arp8-module and identify it as an allosteric sensor of linker DNA. Strikingly, the Ino80^{HSA} adopts a segmented conformation comprising three helices that bind to the barbed ends of Arp4, N-actin and Arp8. The Arp8-module binds extranucleosomal DNA, and we identified a conserved positively charged patch on the solvent accessible site of the Ino80^{HSA} as responsible for DNA binding. Structure based mutagenesis showed that binding of extranucleosomal DNA by Ino80^{HSA} is critical for INO80 nucleosome sliding, but not for INO80 nucleosome binding and ATP hydrolysis. Thus, linker DNA sensing by the Arp8-module drives remodelling by INO80 via coupling motor activity to nucleosome repositioning.

Results

Crystal structure of the 180 kDa Ino80^{HSA}-Arp4-N-actin-Arp8-module

To gain molecular and functional insights into the evolutionarily conserved Arp8-module of INO80, we determined its crystal structure **(Fig. 1)**. N-actin, Arp4, Arp8 (residues 255-881, excluding the non-conserved N-terminal region[27]) and Ino80^{HSA} (residues 461-598) from *S. cerevisiae* were produced in insect cells as a stoichiometric 180 kDa complex **(Supplementary Fig. 1a).** Crystallization initially failed, most likely due to structural flexibility. In a recent study, N-actin adopted a nucleotide free state[9] bound to Arp4 and Swr1^{HSA}, whereas early biochemical analysis of N-actin in the human BAF-complex[28] as well as our structural analysis using a cameloid nanobody (see below) indicated ATP binding of N-actin. Consequently, we sought to limit the structural heterogeneity of the Arp8-module by using latrunculin A (LatA), a small molecule sea sponge toxin that inhibits nucleotide exchange of monomeric actin[29]. Addition of LatA yielded crystals of the complex diffracting to 4 Å, and the

structure was determined by molecular replacement (see **Table 1** for refinement and model statistics).

Figure 1 shows the elongated architecture of the Arp8-module. Ino80^{HSA} forms a markedly segmented α-helix with helical elements α1', α1'' and α2, spanning a distance of in total 120 Å **(Fig. 1a)**. All three actin-fold proteins bind via their barbed ends to the different HSA helical elements in a similar and serial fashion, while pointed ends remain accessible (**Fig. 1c**). From Ino80^{HSA}'s N- to C-terminus, the order of binding is Arp4 (to α1'), N-actin (to α1'') and Arp8 (to α2). The segmentation of the HSA helix enables N-actin to form multiple contacts to both Arp4 and Arp8. Arp4 engages N-actin in a "front-to-back" orientation in contrast to the classical fibrous (F) actin "front to front" interaction[30] (**Fig. 1d**). However, the staggered packing of their subdomains (SDs) as well as local contacts resemble lateral interactions of two F-actin subunits in a filament. In contrast, Arp8 packs against the lateral face of N-actin opposite of Arp4 by using a fundamentally different "side-to-front" type of interaction, unlike any other seen so far between actin-fold proteins. Interestingly, we observed unambiguous density for ATP in the nucleotide binding pocket of Arp4 and N-actin, whereas Arp8 remains nucleotide free (**Supplementary Fig. 1b and c**). Constitutive ATP binding by Arp4 is consistent with our previous observations suggesting that Arp4 is catalytically inactive[26]. However, N-actin may still retain its activity as part of chromatin remodellers[28] and was captured here in its ATP state by LatA. Of note, ATP must have been co-purified with the complex from the cellular environment, as we did not add any nucleotides and LatA was added after purification.

N-actin and Arp4: a conserved heterodimer in distinct chromatin complexes

Arp4-N-actin within the Arp8-module has an overall configuration similar to Arp4-N-actin bound to Swr1^{HSA} (Arp4-N-actin heterodimers align with an backbone RMSD of 0.96 Å (number of aligned residues (N_{align}) 727; using SSM[31] in COOT[32] see methods for further information) and Arp7-Arp9 bound to Snf2^{HSA} [9,10] suggesting that the Arp4-N-actin heterodimer is a structurally conserved module within the INO80 and SWI/SNF families.

To probe the Arp4-N-actin heterodimer in its native environment, we capitalized on a nanobody (denoted NactNB) that we generated from an alpaca immunized with the endogenous *S. cerevisiae* INO80 complex. Nanobodies emerged as a valuable technology to reveal physiologically important states of cellular key components[33,34]. NactNB is highly selective for the endogenous Arp4-N-actin heterodimer. Affinity enrichment mass spectrometry (AE MS) of yeast whole cell lysate using NactNB showed all 35 subunits of chromatin-associated yeast complexes containing the Arp4-N-actin heterodimer (INO80,

SWR1 and NuA4) (**Fig. 2a, b**) suggesting that NactNB recognises a solvent exposed and conserved feature in all of these complexes. To reveal this binding epitope, we determined crystal structures of the Arp4-N-actin-NactNB ternary complex (**Fig. 2c, Supplementary Fig. 2a and Table 1**). NactNB binds into a crevice jointly formed by the pointed ends of the two actin-folds opposite of the Ino80^{HSA} binding site (**Supplementary Fig. 2b**). Satisfyingly, NactNB thus recognises the same staggered configuration of N-actin and Arp4, as present in the structure of the Arp8-module (Arp4-N-actin heterodimers align with an RMSD of 0.68 Å and N_{align} 753) and in complex with Swr1^{HSA} (Arp4-N-actin heterodimers align with an RMSD of 0.96 Å and N_{align} 724). Moreover, residual density in the nucleotide binding pocket in absence of added nucleotide as well as co-crystallization with ATP showed that NactNB recognises the ATP state of N-actin (**Supplementary Fig. 2c, d and e**). NactNB detects the relative orientation of the two N-actin lobes and inserts Arg104 in between SD2 and SD4 where it makes hydrogen bonds to the ATP-bound conformation of Ser14 and Asp157 of the phosphate binding loop P1 and P2, respectively, as well as Glu72 of the ATP sensor loop (**Fig. 2c**).

Taken together, our data provide direct evidence for a conserved configuration of the Arp4-N-actin heterodimer in the complete endogenous INO80, SWR1 and NuA4 complexes and suggest that N-actin can adopt an ATP bound state in its native environment, as previously also suggested for the human BAF complex[28]. The conserved nature of the Arp4-N-actin heterodimer may point towards a common yet so far unknown functional role of this module in distinct chromatin complexes.

Arp8 recruits Arp4-N-actin to a segmented "two plug" scaffold of Ino80^{HSA}

Deletion of Arp8 resulted in partially assembled INO80 lacking also Arp4 and N-actin[15,16]. It rendered yeast cells highly sensitive to metabolic and genotoxic stress[15]. A similar phenotype was observed upon partial removal of the Ino80^{HSA} and post-HSA domain (Ino80$^{post\text{-}HSA}$) (residues 531-598)[23]. The structure of the Arp8-module provides a framework for rationalizing the importance of Arp8 and the Ino80^{HSA} for recruitment of the Arp4-N-actin heterodimer to the INO80 complex (**Fig. 3a**). Arp8 directly engages N-actin through contacts between SD1 and SD2 of Arp8 with SD3 and SD4 of N-actin. In addition, we identify a function for long insertion element I-3a of Arp8. I-3a covers the lateral surface of the Arp8 actin-fold and forms thereby a latch that consolidates the interaction with N-actin. Overall, this bipartite interaction of Arp8 recognises a 1392 Å^2 large area of the N-actin lateral face opposite of Arp4 and thus specifically helps to recruit and retain the interaction of Arp4-N-actin with Ino80^{HSA}.

Previous models proposed that N-actin and Arps are recruited to chromatin remodellers by a long, continuous HSA helix that provides a binding platform for barbed ends of actin-fold proteins[9,10]. While the general helical structure and serial binding of Arp4 and N-actin barbed ends is consistent with this model, Ino80^{HSA} adopts a distinct segmented structure (**Fig. 3b, Supplementary Fig. 3a, b**). The N-terminal helix α1 (residues 472-518), bound to the barbed ends of Arp4 and N-actin, has a pronounced kink at position 483-485 that divides it into segments α1' and α1''. The C-terminal helix α2 (residues 522-557) forms the third segment, bound to the barbed end of Arp8. We identified two hydrophobic residue clusters (Plug 1 and Plug 2) that define the register and contain each an anchoring tryptophan residue. A structural shift resulting from segmentation of α1 enables Plug 1 (Ile476, Trp477 and Met480) to insert into a hydrophobic pocket of the barbed end of Arp4 (**Fig. 3b, Supplementary Fig. 3c**), while well-defined loop L1 in between α1'' and α2 shifts α2 enabling insertion of Plug 2 (Met535, Phe542 and Trp543) into a hydrophobic pocket of the barbed end of Arp8. The latter interaction appears to be critical not only for recruitment of Arp8 but also of the Arp4-N-actin dimer. The previously reported partial removal of the Ino80^{HSA} and Ino80$^{post\text{-}HSA}$ includes Plug2 of α2 and leads to loss of the entire Arp8-module *in vivo*[23], although the Arp4-N-actin binding site of the Ino80^{HSA} is still intact. The distance between the two hydrophobic plugs in conjunction with the asymmetry of Ino80^{HSA} segmentation matches the unique arrangement of actin-folds within the sandwich-like structure of Arp4, N-actin and Arp8. In addition, loop L1 and the resulting translational and rotational shift of α2 enables formation of the extensive contacts between N-actin and Arp8 that would not be possible for a continuous HSA helix. Thus, our structure shows how Arp8 specifies recruitment of the Arp4-N-actin heterodimer to the segmented, "two plug" scaffold of the helical Ino80^{HSA}.

Ino80^{HSA} of the Arp8-module binds extranucleosomal DNA

Our recent cryoEM study of the INO80^{Core}-nucleosome complex revealed density of the Arp8-module adjacent to the well resolved nucleosomal DNA entry site, where the Ino80 Snf2-motor domain pumps DNA into the nucleosome[24] (**Fig. 4a**). To test for binding of the Arp8-module to nucleosomal and extranucleosomal DNA, we performed electro mobility shift assays (EMSAs) where nucleosomes with (0N80) and without (0N0) 80 bp extranucleosomal DNA on one side compete for binding the Arp8-module (**Fig. 4b**). In such competition assays, the Arp8-module showed a clear binding preference for the 0N80 over the 0N0 nucleosome, showing that the Arp8-module binds extranucleosomal DNA.

Combination of the INO80^{Core}-nucleosome complex cryoEM structure[24] and the Arp8-module crystal structure leads directly to a structural model of how the Arp8-module might be located

at extranucleosomal DNA, as discussed further below. In this model, Ino80^{HSA} mediates direct binding of extranucleosomal DNA along the barbed ends of Arp8, N-actin and Arp4. In isolation, neither actin and Arp4 nor human Arp8 bind dsDNA with considerable affinity[26], while the Ino80^{HSA} was proposed from sequence analysis to be part of a dsDNA binding domain of Ino80[35]. Having the crystal structure for Ino80^{HSA}, we noticed a set of highly conserved, solvent accessible lysine and arginine residues that may account for binding of extranucleosomal DNA (**Fig. 3b and Supplementary Fig. 4a**). To test this hypothesis, we mutated several of these lysine and arginine residues in the Ino80^{HSA} α2 helix to glutamines (HSAα2). We observed lower expression yields of the mutated minimal Arp8-module, indicating perhaps destabilizing effects of the mutations by lowering the helix propensity of Ino80^{HSA}. However, using complex-stabilizing NactNB for purification provided sufficient quantities of stable material for DNA binding studies (**Supplementary Fig. 4b**). Fluorescence anisotropy analysis upon binding of generic 40 bp dsDNA and competition EMSAs with 0N0 and 0N80 nucleosomes showed that binding of NactNB at the pointed end of N-actin only slightly reduced dsDNA binding (around 2-fold, **Supplementary Fig. 4c, d**). So we used NactNB to rule out that any loss of DNA binding is induced by weakening of the complex. Importantly, the α2-mutations substantially reduced binding of the Arp8-module both to dsDNA (**Supplementary Fig. 4c**) and nucleosomes (**Fig. 4b** and **Supplementary Fig. 4d**). Thus, we conclude that the positively charged HSA domain of Ino80 provides a binding site for extranucleosomal DNA.

Arp8-module is important for nucleosome sliding and genome-wide nucleosome positioning

To assess the mechanistic impact of DNA binding by the Arp8-module on nucleosome remodelling by INO80 (**Fig. 4**), we mutated Ino80^{HSA} in the context of the entire INO80 complex (**Supplementary Fig. 4e**). Parallel to this study, we established an insect cell co-expression approach for expression and purification of the entire *S. cerevisiae* 15-subunit INO80 complex. Such recombinant INO80 retains the activity of the endogenous complex, but is fully amendable to site directed mutagenesis (to be published elsewhere by: Krietenstein Nils, Oberbeckmann Elisa, Niebauer Vanessa, Schall Kevin, Schwarz Marianne, Moldt Manuela, Tobias Straub, Korber Philipp, Hopfner Karl-Peter & Eustermann Sebastian). Using this system, we were able to purify stable INO80 complexes with WT-like stoichiometry of all subunits, also if full length Ino80 with mutated HSA was co-expressed together with all other 14 subunits of INO80 (**Supplementary Fig. 4f**). EMSAs with 0N80 nucleosomes showed homogenous complex formation at similar concentrations for wild type (WT) as well as for mutant INO80 (**Supplementary Fig. 4g**). This was in contrast to the decreased binding affinity of Ino80-HSAα2 in context of the isolated Arp8-module and suggests that

binding of the entire complex to nucleosomes is mostly dominated by subunits other than the minimal Arp8-module, e.g., subunits of the INO80 core that interact directly with the nucleosome, or other DNA binding subunits like the Nhp10-module. Of note, Arp8 in the recombinant 15 subunit INO80 complex contains the full N-terminal tail in contrast to the construct used for crystallization. Although the N-terminal region of Arp8 is not well conserved among species it might additionally contribute to nucleosome interactions, DNA binding or complex stability.

However, despite retaining high-affinity nucleosome interactions, mutations of the Ino80^{HSA} domain markedly affected dsDNA-stimulated ATPase and ATP-dependent nucleosome sliding activity of INO80. ATP hydrolysis by WT INO80 is robustly stimulated upon addition of dsDNA or 0N80 nucleosomes **(Fig. 4c)**. Mutations of the helix α1 or α2 of Ino80^{HSA} impaired ATPase stimulation by dsDNA, while the same mutants showed similar or even moderately faster ATP hydrolysis rates than WT complex if stimulated by 0N80 nucleosomes. Despite this similar or increased ATP turnover, HSA mutations substantially decreased INO80's activity to slide 0N80 nucleosomes towards the center of a 225 bp DNA substrate (**Fig. 4d**). Mutations of either helix α1 or α2 reduced nucleosome centering to residual levels, while mutations targeting both helices abrogated this remodelling activity completely.

Given this mechanistic impact on sliding activity in a specialized mononucleosome context, we asked whether Ino80^{HSA} is also more generally important to mobilize and position nucleosomes across the whole yeast genome. To this end, we employed a genome-wide reconstitution approach, where it was shown previously that purified INO80 on its own is able to properly position +1 nucleosomes on a genomic plasmid library[2]. A similarly direct analysis of INO80's nucleosome positioning activity would be inherently difficult *in vivo* given the complex interplay between different remodeller families as well as other factors such as the transcription and replication machinery[2,36]. MNase-seq was used as read out for nucleosome positions across the genomic plasmid library before and after incubation with INO80 and ATP. In contrast to the strongly decreased sliding activity with a mononucleosomal substrate based on the "Widom 601" sequence (**Fig. 4d**), INO80 mutations targeting HSA helix α1 or α2 individually did not compromise average patterns of genomic +1 nucleosome positioning (**Fig. 4e and f**). This finding is intriguing as it suggests that nucleosomes on genomic DNA in plasmids may be a less demanding substrate for translocation and positioning than a "Widom 601" mononucleosome. This could be, for example, due to the presence of multiple nucleosomes, extranucleosomal DNA on both sides, the absence of DNA ends, or due to lower intrinsic nucleosome stabilities. The former three possibilities seemed unlikely as remodeling an internal nucleosome within an array of 601 sequences separated by 50 bp

extranucleosomal DNA[37] was also strongly impaired by mutations targeting either HSA helix individually (**Supplementary Fig. 4h and i**). Importantly, however, mutation of both HSA helices α1 and α2 at the same time abolished INO80 nucleosome remodelling on all tested substrates including genome-wide nucleosome positioning **(Fig. 4d, e, f and supplementary Fig. 4 i).**

Taken together, our biochemical results identify a critical role for binding extranucleosomal DNA by the Arp8-module in coupling the energy derived from ATP hydrolysis to productive nucleosome sliding by INO80. Such chemo-mechanical coupling may be particularly important to mobilize nucleosomes in the context of sequences that strongly bind the histone octamer, such as the 601 sequence. Positioning of +1 nucleosomes guided by promoter sequences is likely to involve also other parts of the INO80 complex such as the Nhp10-module.

A Structural Model of the INO80-nucleosome complex including its Arp8-module

Combination of the 4.3 Å cryoEM structure of the *C. thermophilum* INO80^{Core}-nucleosome complex[24] with the 4 Å *S. cerevisiae* crystal structure of the 180 kDa Arp8-module leads directly to a composite model of the evolutionarily conserved INO80$^{Core+Arp}$complex bound to a nucleosome with a molecular weight of approximately 1 MDa (**Fig. 5a**). The two structures can be joined in silico by the highly conserved post-HSA and HSA domains of Ino80: helix α2 of the HSA domain crystal structure needs to be extended by only 35 C-terminal residues to include the post-HSA helix that is present in the cryoEM structure[24]. This structural model is consistent with the mapping of INO80 subunits onto nucleosomal substrates *in vivo*[38], *in vitro* (unpublished data by: Sandipan Brahma, Mzwanele Ngubo, Somnath Paul, Maheshi Udugama, and Blaine Bartholomew) and with our previous cryoEM data[24] as it places the Arp8-module into the large unassigned density patch (**Fig. 4a and 5b**) and at the same time maintains a continuous HSA and post-HSA helical structure. In particular, we observed cryoEM density for an extended post-HSA-HSA helix pointing from the N-terminal lobe of the Snf2-type ATPase domain at SHL-6 towards entry DNA at SHL-8[24]. Moreover, the elongated architecture of the Arp8-module accommodates approximately 40 bp of extranucleosomal entry DNA and fits thereby into the low-resolution reconstruction obtained for the entire 11-subunit INO80$^{Core+Arp}$-nucleosome complex[24]. The 120 Å HSA domain is positioned along the dsDNA with conserved arginine and lysine residues contacting the phosphate backbone as probed by our biochemical experiments described above. Helix α2 contacts the DNA around SHL-8 while the N-terminal helix α1 reaches SHL-10 to -11. Consequently, Arp8 resides on the extranucleosomal DNA proximal to the Snf2-type ATPase of Ino80, while the Arp4-N-actin heterodimer binds in a distal position. The model is therefore consistent with promoter

binding of Arp8 proximal to the +1 nucleosome *in vivo*[38] as well as crosslinking of Arp4-N-actin heterodimer to extranucleosomal DNA at position -110 nt (SHL-11) *in vitro* (unpublished data by: Sandipan Brahma, Mzwanele Ngubo, Somnath Paul, Maheshi Udugama, and Blaine Bartholomew). However, given the flexibility of the Arp8-module in the cryo-EM reconstructions, we do not rule out the presence of other conformations and positions of this module during the functional cycle of INO80 in nucleosome remodelling.

Discussion

Here we provide a structure and function for the enigmatic, evolutionarily conserved actin-fold subunits Arp4, N-actin and Arp8 in the INO80 chromatin remodeller. We show that the three actin-fold proteins in complex with the Ino80^{HSA} form an extended structural element that recognises extranucleosomal, linker DNA, a critical feature of INO80 mechanism and function.

INO80 is a highly processive chromatin remodeller[22,39] and we recently proposed a mechanism by which INO80 core subunits function as a macromolecular ratchet[24]: minor groove tracking by the Ino80 Snf2-type ATPase motor at SHL-6 pumps DNA in multiple 1-2 bp steps against a grip formed by Arp5-Ies6 at SHL-2/-3 until DNA propagates around the histone octamer and translocates nucleosomes by a large step size. 10-20 bp translocation steps are indeed observed[22,40], and a kinetic model has been proposed describing the dependency of INO80 on extranucleosomal DNA[22]: the activity of the ATPase motor does not result in efficient DNA translocation unless more than 40 bp of entry DNA are available; the pumped DNA might otherwise collapse backwards[22]. Intriguingly, the footprint of the Arp8-module matches this limiting length of 40 bp DNA (**Fig. 5a**). If less than 40 bp extranucleosomal, linker DNA are available, pumping an additional 10-20 bp DNA into the nucleosome would substantially shorten the entry DNA beyond this limit, i.e., pull away the DNA and thereby abrogate the contacts between DNA and the distal part of the HSA domain, where the Arp4-N-actin heterodimer binds. Consequently, this scenario recapitulates the impact of HSA mutations that also lead to loss of extranucleosomal DNA binding and reduce nucleosome sliding to residual levels, most likely caused by “back-slippage” of DNA. By combining our structural and biochemical data with previous kinetic insights[22], we thus propose that the Arp8-module within INO80 functions as sensor of extranucleosomal DNA, mechanistically coupling ATP-dependent DNA pumping to processive nucleosome translocation.

Biochemical and genetic evidence for the SWI/SNF chromatin remodeller family suggests that the yeast Arp7-Arp9-module of RSC has a role similar to that proposed here for the

Arp8-module of INO80, as it also couples ATP-dependent DNA translocation of the Snf2-type Sth1 motor domain to nucleosome remodelling such as translocation and ejection[41]. It was proposed that the post-HSA domain of Sth1 acts as a "throttle" controlling ATPase activity[41]. Indeed, our structural study shows that the post-HSA domain interacts with the N-lobe of Ino80[24] in a homologous manner as previously observed for Snf2[42] and Sth1[43]. A structure based alignment reveals that the highly conserved (Q)TELY motif[6] of the Ino80$^{post\text{-}HSA}$ domain is related to the QTXX[F/Y] motif of Snf2 (**Fig. 6a and b**), while the interaction with protrusion-I provides, together with brace helix-I, a key allosteric site for controlling DNA groove tracking by the ATPase motor[41,42]. Despite recent progress[10,41,43,44], it is still elusive how the Arp7-Arp9-module of RSC functions at a molecular level. It was suggested that the module folds back onto the Sth1 motor domain acting as "clutch" to promote nucleosome remodelling. While INO80 might adopt a closed conformation in absence of a nucleosome[45,46], our structural and biochemical data suggest an extended conformation of the Arp8-module which enables extranucleosomal DNA binding. The interplay between the HSA and post-HSA domains may thus link sensing of extranucleosomal DNA to allosteric control of the Snf2-like motor domain of Ino80.

Sensing of linker DNA is a hallmark of chromatin remodellers since it provides mechano-chemical means to conduct higher order remodelling reactions such as spacing and phasing of nucleosomes in genic arrays[4,47]. Future studies will use the mechanistic insights discovered here as a framework to dissect such functions and will show whether they may provide unifying principles for regulation of the INO80 and SWI/SNF families of multi-subunit chromatin remodellers.

Accession codes

Coordinates and structure factors have been deposited in the Protein Data Bank under PDB ID accession codes **5NBM** for the N-actin(ATP)-Arp4-NactNB, **5NBL** for the N-actin(apo)-Arp4-NactNB and **5NBN** for the Ino80^{HSA}-Arp4-N-Actin-Arp8-module structures.
Data of the genome-wide nucleosome positioning experiments reported in this paper have been deposited on the NCBI Gene Expression Omnibus (accession number **GSE113401**).

Acknowledgements

We are grateful to Manuela Moldt for excellent technical support, Jens Michaelis, Gregor Witte, Katja Lames and Robert Byrne for discussion and technical help. We thank the Max-Planck Crystallization Facility (Martinsried, Germany), the staff of the Swiss-Light-Source (SLS, Villingen, Switzerland) and the European-Molecular-Biology-Laboratory/Deutsches-Electronen-Synchotron (EMBL/DESY, Hamburg, Germany) for support and measurement time. We thank Stefan Krebs and Helmut Blum at the Laboratory of Functional Genome Analysis (LAFUGA, Gene Center, LMU Munich) for high-througput sequencing. We thank Tobias Straub (Bioinformatics Core Unit, Biomedical Center, LMU Munich) for advice on bioinformatics. This work is supported by the Deutsche Forschungsgemeinschaft CRC1064 (to K.-P.H and P.K.) and the European Research Council (ERC Advanced Grant ATMMACHINE), the Gottfried-Wilhelm-Leibniz Prize and the Center for Integrated Protein Sciences Munich (CIPSM) to K.-P.H. K.R.K. is supported by GRK1721. S.E. acknowledges an EMBO long term fellowship, V.N., K.S. and M.S. acknowledge funding by Quantitative Biosciences Munich (QBM).

Author contribution

K.R.K and S.E. determined the structures and built atomic models. K.R.K. prepared samples for crystallization and performed biochemical analysis of the Arp8-module. S.E., A.T., M.S. and A.B identified the Arp4-N-actin binding nanobody and performed its initial characterization. K.R.K. and G.S. performed AE MS analysis. S.E. and K.P.H devised with a contribution of M.S. preparation and characterization of recombinant INO80 complex. V.N. prepared mutant complexes and performed their biochemical analysis. V.N. and E.O. performed and analyzed genome-wide remodelling assays under supervision by P.K. K.S. prepared nucleosomes. S.E. and K.-P.H. designed the overall study, analyzed the results and wrote the paper with contributions from K.R.K., V.N., E.O. and P.K.

Competing financial interests statement

The authors declare no competing financial interest.

References

1. Jiang, C. & Pugh, B.F. Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* **10**, 161-72 (2009).
2. Krietenstein, N. et al. Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell* **167**, 709-721 e12 (2016).
3. Hopfner, K.P., Gerhold, C.B., Lakomek, K. & Wollmann, P. Swi2/Snf2 remodelers: hybrid views on hybrid molecular machines. *Curr Opin Struct Biol* **22**, 225-33 (2012).
4. 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).
5. Dion, V., Shimada, K. & Gasser, S.M. Actin-related proteins in the nucleus: life beyond chromatin remodelers. *Curr Opin Cell Biol* **22**, 383-91 (2010).
6. Shen, X., Mizuguchi, G., Hamiche, A. & Wu, C. A chromatin remodelling complex involved in transcription and DNA processing. *Nature* **406**, 541-4 (2000).
7. Peterson, C.L., Zhao, Y. & Chait, B.T. Subunits of the yeast SWI/SNF complex are members of the actin-related protein (ARP) family. *J Biol Chem* **273**, 23641-4 (1998).
8. Cairns, B.R., Erdjument-Bromage, H., Tempst, P., Winston, F. & Kornberg, R.D. Two actin-related proteins are shared functional components of the chromatin-remodeling complexes RSC and SWI/SNF. *Mol Cell* **2**, 639-51 (1998).
9. Cao, T. et al. Crystal structure of a nuclear actin ternary complex. *Proc Natl Acad Sci U S A* **113**, 8985-90 (2016).
10. Schubert, H.L. et al. Structure of an actin-related subcomplex of the SWI/SNF chromatin remodeler. *Proc Natl Acad Sci U S A* **110**, 3345-50 (2013).
11. Szerlong, H. et al. The HSA domain binds nuclear actin-related proteins to regulate chromatin-remodeling ATPases. *Nat Struct Mol Biol* **15**, 469-76 (2008).
12. Meagher, R.B., Kandasamy, M.K., Smith, A.P. & McKinney, E.C. Nuclear actin-related proteins at the core of epigenetic control. *Plant Signal Behav* **5**, 518-22 (2010).
13. Son, E.Y. & Crabtree, G.R. The role of BAF (mSWI/SNF) complexes in mammalian neural development. *Am J Med Genet C Semin Med Genet* **166C**, 333-49 (2014).
14. Hodges, C., Kirkland, J.G. & Crabtree, G.R. The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer. *Cold Spring Harb Perspect Med* **6**(2016).
15. Shen, X., Ranallo, R., Choi, E. & Wu, C. Involvement of actin-related proteins in ATP-dependent chromatin remodeling. *Mol Cell* **12**, 147-55 (2003).
16. Tosi, A. et al. Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* **154**, 1207-19 (2013).
17. Gerhold, C.B. & Gasser, S.M. INO80 and SWR complexes: relating structure to function in chromatin remodeling. *Trends Cell Biol* **24**, 619-31 (2014).
18. Papamichos-Chronakis, M., Watanabe, S., Rando, O.J. & Peterson, C.L. Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* **144**, 200-13 (2011).
19. Udugama, M., Sabri, A. & Bartholomew, B. The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. *Mol Cell Biol* **31**, 662-73 (2011).
20. Chen, L. et al. Subunit organization of the human INO80 chromatin remodeling complex: an evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling. *J Biol Chem* **286**, 11283-9 (2011).
21. Jonsson, Z.O., Jha, S., Wohlschlegel, J.A. & Dutta, A. Rvb1p/Rvb2p recruit Arp5p and assemble a functional Ino80 chromatin remodeling complex. *Mol Cell* **16**, 465-77 (2004).

22. Zhou, C.Y. et al. The Yeast INO80 Complex Operates as a Tunable DNA Length-Sensitive Switch to Regulate Nucleosome Sliding. *Mol Cell* **69**, 677-688 e9 (2018).
23. Kapoor, P., Chen, M., Winkler, D.D., Luger, K. & Shen, X. Evidence for monomeric actin function in INO80 chromatin remodeling. *Nat Struct Mol Biol* **20**, 426-32 (2013).
24. Eustermann, S. et al. Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* (2018).
25. Ayala, R. et al. Structure and regulation of the human INO80–nucleosome complex. *Nature* (2018).
26. Gerhold, C.B. et al. Structure of Actin-related protein 8 and its contribution to nucleosome binding. *Nucleic Acids Res* **40**, 11036-46 (2012).
27. Saravanan, M. et al. Interactions between the nucleosome histone core and Arp8 in the INO80 chromatin remodeling complex. *Proc Natl Acad Sci U S A* **109**, 20883-8 (2012).
28. Zhao, K. et al. Rapid and phosphoinositol-dependent binding of the SWI/SNF-like BAF complex to chromatin after T lymphocyte receptor signaling. *Cell* **95**, 625-36 (1998).
29. Dominguez, R. & Holmes, K.C. Actin structure and function. *Annu Rev Biophys* **40**, 169-86 (2011).
30. von der Ecken, J. et al. Structure of the F-actin-tropomyosin complex. *Nature* **519**, 114-7 (2015).
31. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256-68 (2004).
32. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
33. Huang, W. et al. Structural insights into micro-opioid receptor activation. *Nature* **524**, 315-21 (2015).
34. Rasmussen, S.G. et al. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549-55 (2011).
35. Bakshi, R., Prakash, T., Dash, D. & Brahmachari, V. In silico characterization of the INO80 subfamily of SWI2/SNF2 chromatin remodeling proteins. *Biochem Biophys Res Commun* **320**, 197-204 (2004).
36. Yen, K., Vinayachandran, V., Batta, K., Koerber, R.T. & Pugh, B.F. Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell* **149**, 1461-73 (2012).
37. Mueller-Planitz, F., Klinker, H., Ludwigsen, J. & Becker, P.B. The ATPase domain of ISWI is an autonomous nucleosome remodeling machine. *Nat Struct Mol Biol* **20**, 82-9 (2013).
38. Yen, K., Vinayachandran, V. & Pugh, B.F. SWR-C and INO80 chromatin remodelers recognize nucleosome-free regions near +1 nucleosomes. *Cell* **154**, 1246-56 (2013).
39. Schwarz, M. et al. Single-molecule nucleosome remodeling by INO80 and effects of histone tails. *FEBS Letters* **592**, 318-331 (2018).
40. Brahma, S. et al. INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers. *Nature Communications* **8**, 15616 (2017).
41. Clapier, C.R. et al. Regulation of DNA Translocation Efficiency within the Chromatin Remodeler RSC/Sth1 Potentiates Nucleosome Sliding and Ejection. *Mol Cell* **62**, 453-61 (2016).
42. Liu, X., Li, M., Xia, X., Li, X. & Chen, Z. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* **544**, 440 (2017).

43. Turegun, B., Baker, R.W., Leschziner, A.E. & Dominguez, R. Actin-related proteins regulate the RSC chromatin remodeler by weakening intramolecular interactions of the Sth1 ATPase. *Communications Biology* **1**, 1 (2018).
44. Turegun, B., Kast, D.J. & Dominguez, R. Subunit Rtt102 controls the conformation of the Arp7/9 heterodimer and its interactions with nucleotide and the catalytic subunit of SWI/SNF remodelers. *Journal of Biological Chemistry* (2013).
45. Aramayo, R.J. et al. Cryo-EM structures of the human INO80 chromatin-remodeling complex. *Nat Struct Mol Biol* **25**, 37-44 (2018).
46. Watanabe, S. et al. Structural analyses of the chromatin remodelling enzymes INO80-C and SWR-C. *Nat Commun* **6**, 7108 (2015).
47. Yamada, K. et al. Structure and mechanism of the chromatin remodelling factor ISW1a. *Nature* **472**, 448-53 (2011).
48. Lin, C.-L. et al. Functional characterization and architecture of recombinant yeast SWR1 histone exchange complex. *Nucleic Acids Research* **45**, 7249-7260 (2017).

Figures and figure legends

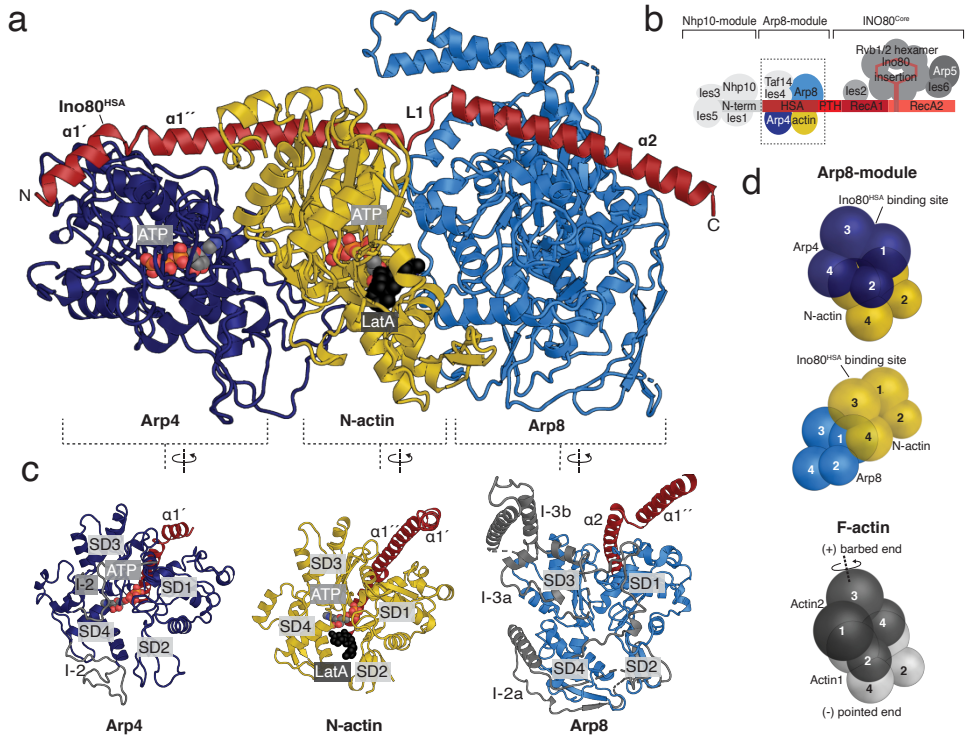


Figure 1. Crystal structure of the 180 kDa Arp4-N-Actin-Arp8-Ino80^{HSA} module.
(a) Crystal structure of the INO80 Arp8-module comprising Arp4, N-actin, Arp8 and Ino80^{HSA}. (**b**) Schematic overview of the *S. cerevisiae* INO80 complex illustrating its modular architecture. (**c**) Front views of the actin-fold proteins Arp4, N-actin and Arp8. The Ino80^{HSA} binds to the barbed end of each of the actin folds. Arp4 and N-actin are ATP-bound (coloured spheres), whereas Arp8 is nucleotide-free. Latrunculin A (LatA, black spheres) is bound next to ATP in the N-actin nucleotide binding cleft. Actin fold insertions of Arp4 and Arp8 are shown in grey. (**d**) Arrangement of actin-fold proteins. Schematic actin-folds with the individual subdomains as spheres. Interaction of Arp4 to N-actin and N-actin to Arp8 in the Arp8-module in comparison to two laterally interacting actin molecules in F-actin.

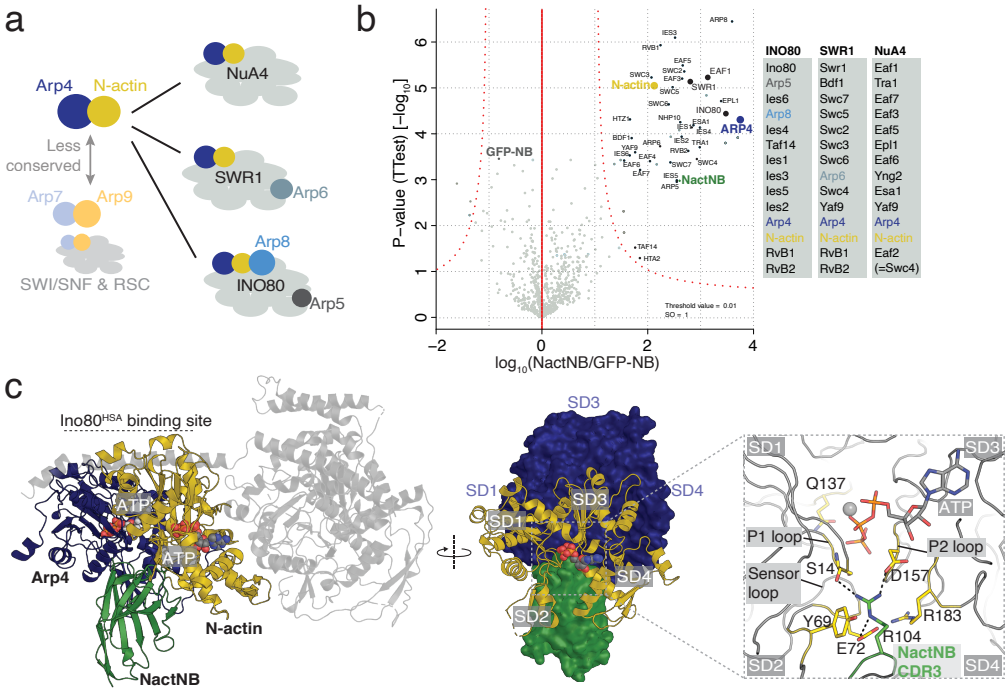


Figure 2. Arp4-N-actin heterodimer: a conserved structural module of chromatin complexes. (**a**) Arp4 and N-actin are conserved core components of all INO80 and SWI/SNF chromatin remodeller families, except for the *S. cerevisiae* SWI/SNF and RSC remodellers, which instead contain the sequence divergent Arp7 and Arp9. INO80 and SWR1 contain in addition Arp5, Arp6 and Arp8. (**b**) NactNB captures endogenous Arp4-N-actin heterodimer. Yeast whole cell extract was subjected to AE MS experiments using NactNB and a GFP-binding nanobody as a control. Experiments were performed in triplicates and a two-sided and two-sampled t-test shows in a volcano plot representation significant enrichment of all 34 subunits of INO80, SWR1 and NuA4 complexes (see methods for further information). (**c**) Structure of the Arp4-N-actin-NactNB complex in two orientations as cartoon and surface representation (left panel: in light grey the Arp8-module structure aligned on the Arp4-N-actin dimer). N-actin and Arp4 are ATP-bound (coloured spheres). Close-up displays that Arg104 of NactNB binds the nucleotide binding cleft of N-actin.

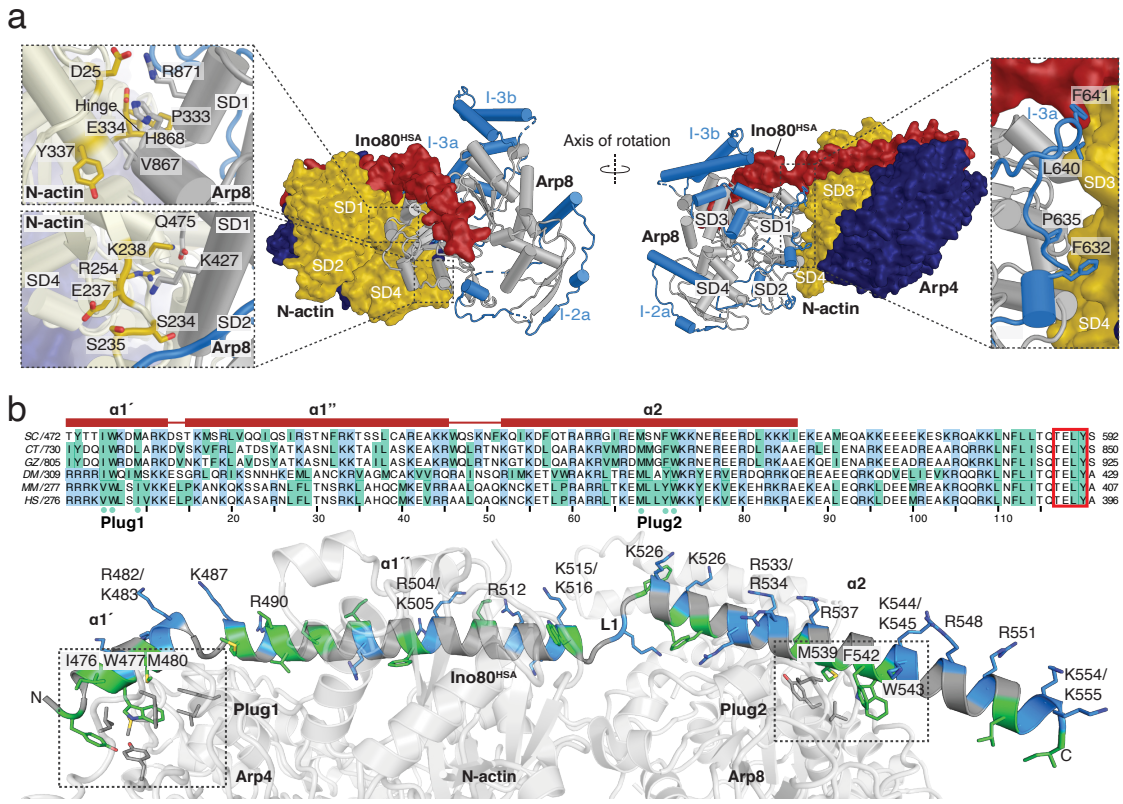


Figure 3. Arp8 recruits Arp4-N-actin to a segmented “two plug” scaffold of Ino80^{HSA}.
(**a**) Cartoon and surface representation of the Arp8-module displaying interaction sites between Arp8 and N-actin. The Arp8 actin core fold is coloured in grey and the insertions in blue. Arp8 contacts N-actin SD3 and 4 via its actin core fold, with SD1 and 2 (close ups in the left panel), and its actin fold insertions 3A (close up in the right panel). (**b)** Sequence alignment of Ino80^{HSA} from different species, with positively charged residues (Arg and Lys) coloured in blue and hydrophobic residues (Ile, Leu, Trp, Val, Phe, Tyr and Met) in green. In the crystal structure visible region is indicated above by a red stroke. The highly conserved TELY motif is highlighted by a red rectangle. Green dots below the sequences emphasize conserved hydrophobic residues, Plug1 and Plug2, that bind to Arp4 and Arp8 respectively. Below, cartoon representation of the Arp8-module. The Ino80^{HSA} domain with hydrophobic residues coloured in green and positive charged residues coloured in blue.

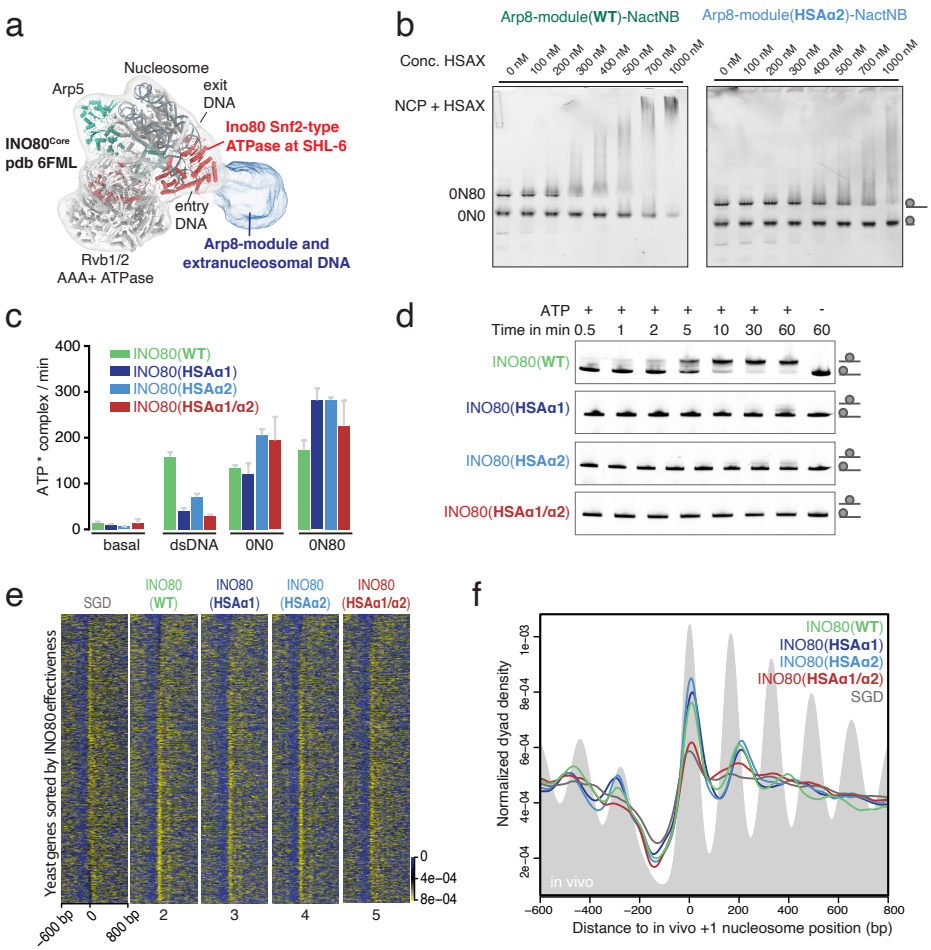


Figure 4. Extranucleosomal DNA binding by the Arp8-module is critical for INO80 nucleosome sliding and genome-wide nucleosome positioning.

(a) CryoEM density of the INO80-nucleosome complex[24], with a structural model for Ino80^{Motor}, Ies2, Arp5, Ies6, and the Rvb1-Rvb2 heterohexamer bound to a nucleosome core particle (NCP). Density next to the nucleosomal DNA entry site could be assigned to the Arp8-module (coloured in blue). **(b)** Competition electro mobility shift assays with two nucleosome species (20 nM each; one with (0N80) and one without (0N0) 80 bp extranucleosomal DNA overhang), showing a clear binding preference of the Arp8-module for 0N80 nucleosomes. Mutations of solvent exposed basic residues on helix α2 of Ino80^{HSA} (HSAα2) decreased 0N80 binding by the Arp8-module. Experiments were performed in triplicates. **(c)** INO80 (27 nM) ATPase activity, basal and stimulated with 223 bp dsDNA (100 nM), 0N0 (100 nM) and 0N80 (50 nM) nucleosomes. Data points and error bars represent the means ± s.d. from 3 independent experiments. **(d)** Time course of ATP-dependent INO80 nucleosome (0N80) sliding on a single mononucleosome substrate (with 18 nM IN80 and 90 nM 0N80). Reaction educt (end positioned nucleosome) and product (center positioned nucleosome) were resolved by NativePAGE. Experiments were performed in triplicates. **(e)** Genome-wide nucleosome positioning by INO80 (18 nM). Heat map displaying colour coded nucleosome dyad density of YCp50 plasmid library yeast genes aligned on the in vivo defined +1 nucleosome dyad (0 bp) position, after sequence intrinsic nucleosome positioning by salt gradient dialysis (SGD), or after additional incubation with indicated WT or mutant INO80 complexes. Rows are sorted according to INO80 effectiveness. **(f)** Composite plots of heat maps shown in (e). Grey background displays *in vivo* nucleosome positioning. Genome-wide nucleosome positioning assays were performed in duplicates. (Uncropped gel images are shown in Supplementary Data Set 1.)

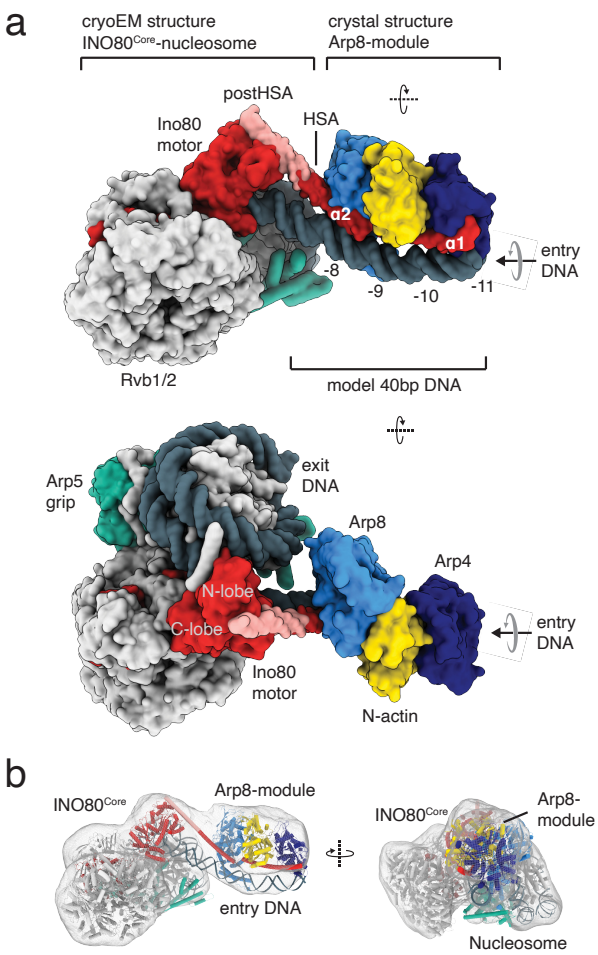


Figure 5. Structural model of the INO80$^{Core+Arp}$-nucleosome complex.
(a) Model of the INO80$^{Core+Arp}$-nucleosome complex (shown as surface representation) based on the INO80^{Core}-nucleosome cryoEM structure[24] and the Arp8-module crystal structure (this study). **(b)** Previously published cryoEM density map of the INO80$^{Core+Arp}$-nucleosome complex[24] with the model of the INO80$^{Core+Arp}$-nucleosome complex fitted into the density.

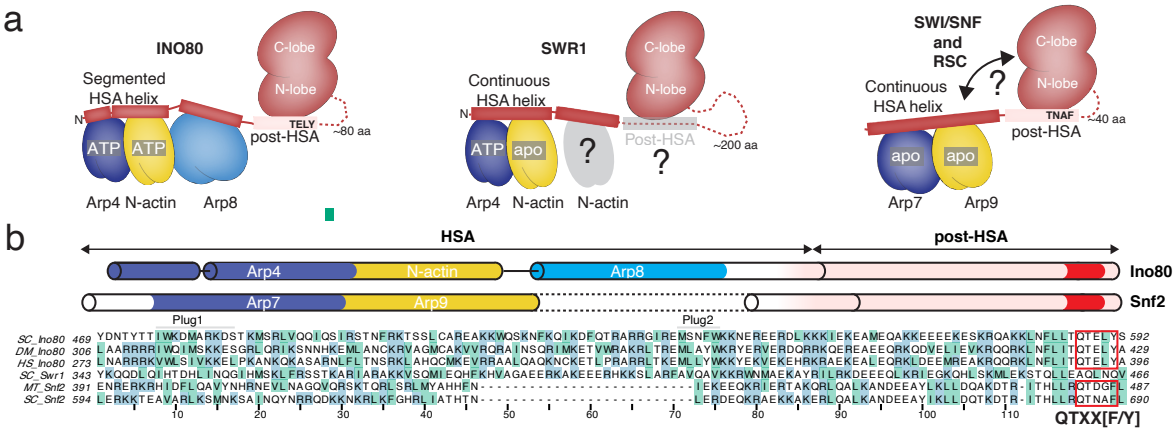


Figure 6. Conserved architecture of N-actin-Arp modules in INO80/SWR1 and SWI/SNF family of chromatin remodeller

(a) Organization of the N-actin-Arp-HSA-modules in INO80, SWR1 and RSC remodeller with respect to the Snf2-type ATPase. The schematic representation is based on our our Ino80$^{Core+Arp}$-nucleosome model (shown in **Fig. 5a**), the structure-based sequence alignment shown in (b) and the crystal structures of Arp4-N-actin-Swr1^{HSA} (PDB ID 5I9E), Arp7-Arp9-Snf2^{HSA} (PDB ID 4I6M), Arp4-N-actin-Arp8-Ino80^{HSA} and Snf2 in complex with a nucleosome (PDB ID 5HZR). Conformation of the respective HSA domain (red) is illustrated by a continuous or segmented helix. The post-HSA of Ino80 and Snf2 (pink) interacts directly with N-lobe of the Snf2-type ATPase (red) and is connected via linker region (dotted line). The nucleotide state of the actin-fold proteins is indicated according to the respective crystal structure. Interestingly, recent biochemical experiments suggested that the Swr1^{HSA} is bound by Arp4 and two N-actin molecules[48]. Our structure based alignment shown in (b) shows that the Arp8 binding site of Ino80^{HSA} corresponds to the second N-actin site in Swr1^{HSA}. **(b)** Structure-based sequence alignment of the HSA (deep red) and post-HSA (light red) domains of Ino80, Snf2 and Swr1 (basic and hydrophobic residues are highlighted in blue and light green, respectively). Binding sites for Arps and N-actin are conserved between the INO80 and SWI/SNF remodeler families. Interestingly, our structure based alignment reveals that the Ino80$^{post-HSA}$ (Q)TELY motif is related to the Snf2$^{post-HSA}$ QTXX[F/Y] motif.

Tables

Table 1 Data collection and refinement statistics.

	NactNB-Arp4-N-actin(ATP) (5NBM)	NactNB-Arp4-N-actin(apo) (5NBL)	Arp4-N-Actin-Arp8-Ino80^{HSA}-module (5NBN)
Data collection			
Space group	P 65	P 65	C 2 2 21
Cell dimensions			
a, *b*, *c* (Å)	190.58 190.58 220.62	191.22 191.22 221.97	172.29 263.91 241.40
α, β, γ (°)	90.00 90.00 120.00	90.00 90.00 120.00	90.00 90.00 90.00
Resolution (Å)	47.73-3.40 (3.50-3.40)[a]	49.43-2.80 (2.90-2.80)	49.40-4.00 (4.10-4.00)
R_{merge}	0.160 (1.081)	0.146 (1.107)	0.236 (1.336)
$I/\sigma(I)$	12.61 (2.19)	12.08 (2.09)	8.71 (1.87)
$CC_{1/2}$	0.996 (0.719)	0.995 (0.617)	0.996 (0.605)
Completeness (%)	100 (100)	100 (100)	100 (100)
Redundancy	6.5 (6.8)	5.9 (5.4)	9.6 (10.0)
Refinement			
Resolution (Å)	47.73-3.40 (3.50-3.40)	49.43-2.80 (2.90-2.80)	49.40-4.00 (4.10-4.00)
No. reflections	62264 (6206)	112476 (11263)	46675 (4625)
R_{work} / R_{free}	0.152 (0.231) / 0.193 (0.281)	0.171 (0.276) / 0.204 (0.316)	0.193 (0.254) / 0.242 (0.288)
No. atoms			
Protein	13949	14000	23029
Ligand/ion	128[b]	64[c]	186[d]
Water	-	119	-
B factors			
Protein	92.30	58.50	121.87
Ligand/ion	85.97	38.80	101.03
Water	-	48.77	-
R.m.s. deviations			
Bond lengths (Å)	0.004	0.004	0.002
Bond angles (°)	0.66	0.70	0.68

Diffraction data from one NactNB-Arp4-N-actin(ATP), one NactNB-Arp4-N-actin(apo) and one Arp4-N-Actin-Arp8-Ino80^{HSA}-module crystal were used to solve the structures. [a]Values in parentheses are for highest-resolution shell. [b]Bound ligands are two ATP and two calcium ions. [c]Bound ligand is one ATP and one calcium ions. [d] Bound ligands are one latrunculin A, two ATP and two calcium ions.

Online methods

Nanobody generation

For generation of the Arp4-N-actin binding nanobody (NactNB) an alpaca was immunized with purified and cross-linked endogenous INO80 complex. INO80 complex for immunization was prepared as earlier described[16]. Alpaca immunization, nanobody library generation and selection of INO80 binding nanobodies were performed as previously published[49] by ChromoTek GmbH (Munich).

Cloning, protein expression & purification

Nanobody (NactNB)

The DNA sequence coding NactNB carrying a C-terminal double Strep-Tag was cloned into a pHEN6 vector upstream of the pelB leader sequence[50]. *Escherichia coli* Rosetta (DE3) cells (Merck Millipore) were transformed with the pHEN6-NactNB vector. Freshly transformed cells were cultured at 37°C in lysogeny broth (LB) containing 100 µg/mL ampicillin. Protein was expressed for 2 h at 22°C after induction with 0.3 mM isopropyl β-d-1-thiogalactopyranoside at an OD_{600} of 0.6. All protein purification steps were performed at 4°C. Cells were harvested by centrifugation and subsequently incubated for 30 min in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 1x protease inhibitor cocktail (PI) (Sigma-Aldrich), 1 mg/mL lysozyme (Carl Roth) and 12.5 units/mL benzonase (Sigma-Aldrich)) for periplasmic lysis. The cell debris was separated by centrifugation. NactNB was purified from the soluble extract via the C-terminal double Strep-Tag using Strep-Tactin Sepharose (IBA) in 50 mM Tris pH 8.0 and 300 mM NaCl. NactNB bound to Strep-Tactin Sepharose was stored at 4°C and used within two days for pull-down assays or eluted with buffer containing 2.5 mM d-Desthiobiotin.

Arp4-N-actin-NactNB complex

S. cerevisiae genes coding for Arp4, Arp8, actin, Taf14, Ies1, Ies2, Ies3, Ies4, Ies5 and Nhp10 were combined in a single pFBDM vector using the MultiBac system[51]. Integration of genes from the pFBDM vector into the baculoviral genome was performed in DH10MultiBac cells, baculovirus generation in *Spodoptera frugiperda* Sf21 insect cells (IPLB-Sf21AE**)** and proteins co-expression in *Trichoplusia ni* High Five insect cells (Invitrogen) according to a published protocol[51]. High Five cells were transfected 1/100 (v/v) with baculovirus. Cells were cultured for 60 h at 27°C until they were harvested by centrifugation. Cells were lysed by sonication in 50 mM Tris pH 8, 300 mM NaCl, 5% glycerol and 1x PI (Sigma-Aldrich). The raw cell lysate was cleared by centrifugation. NactNB bound Strep-Tactin Sepharose was used to isolate the Arp4-N-actin heterodimer from the soluble cell extract. The Arp4-N-actin-NactNB complex was washed with 50 mM Tris pH 8, 300 mM NaCl and 5 % glycerol and eluted with 50 mM Tris pH 8, 300 mM NaCl, 5 % glycerol, and 2.5 mM d-Desthiobiotin. The

ternary complex was further purified by ion-exchange chromatography with a HiTrapQ HP column (GE Healthcare; linear gradient 100 mM to 1 M NaCl) and gel filtration with a Superdex 200 column (GE Healthcare) equilibrated with 20 mM HEPES pH 8 and 200 mM NaCl. Pure protein was concentrated to a final concentration of 16-20 mg/ml, flash frozen in liquid nitrogen and stored at -80˚C.

Arp8-module

Genes encoding *S. cerevisiae* Arp4 and actin were cloned into one pFBDM vector and those coding for *sc* Arp8 (residues 255-881; the non-conserved N-terminal residues 1-254 were deleted[27]) and Ino80^{HSA} (residues 462-598) carrying a C-terminal StrepTag were combined on a second pFBDM vector[51]. Baculoviruses for the respective vectors were generated in Sf21 insect cells as described above. For the co-expression of the four proteins, High Five insect cells (Invitrogen) were co-infected with the two viruses (1/100 (v/v) each), cultivated for 60 h at 27°C and harvested by centrifugation. High Five cells were lysed by sonication in 20 mM HEPES pH 7.8, 100 mM KCl, 2.5 % glycerol and 1x PI (Sigma-Aldrich). The complex was purified from the cleared cell lysate by affinity chromatography using Strep-Tactin Sepharose (IBA), ion exchange chromatography with a HiTrapQ HP column (GE Healthcare; linear gradient 100 mM-800 mM NaCl) and gel filtration with a Superdex 200 column (GE Healthcare) equilibrated with 20 mM HEPES pH 8, 150 mM KCl, 2.5 % glycerol and 1 mM DTT. Peak fractions containing homogenous Arp8-module complex were pooled, concentrated, flash frozen and stored at -80℃.

For the Arp8-module Ino80-HSAα2 mutant (see **Supplementary Table 1** for sequence) a single pACE-BacI vector encoding expression cassettes for *sc* Arp4, actin, Arp8 (residues 255-881) and Ino80-HSAα2 (residues 462-598 + C-terminal StrepTag) was generated by using the latest MultiBac system[24,51]. Generation of the baculovirus, expression in High Five insect cells (Invitrogen) and purification of the WT and the HSAα2 mutant Arp8-module in complex with NactNB was performed in principle as described above. However, prior purification of the respective complex 1 mg of purified NactNB (purification of NactNB is described above) was added directly to 20 mL of cleared insect cell lysate. Further purification followed the procedure described before for the WT Arp8-module.

INO80 complex

Purification of recombinant expressed *S. cerevisiae* INO80 complex from insect cells will be published elsewhere (Unpublished data by: Krietenstein Nils, Oberbeckmann Elisa, Niebauer Vanessa, Schall Kevin, Schwarz Marianne, Moldt Manuela, Korber Philipp, Hopfner Karl-Peter & Eustermann Sebastian). Briefly, two Baculoviruses were generated by MultiBac technology[51] using coding sequences for *S. cerevisiae* Ino80(2x Flag), Rvb1, Rvb2, Arp4, Arp5-His, Arp8, Actin, Taf14, Ies1, Ies2, Ies3, Ies4, Ies5, Ies6 and Nhp10 subcloned into pFBDM vectors. For expression, high five insect cells (Invitrogen) were co-infected with the

two baculoviruses 1/100 (v/v) each. INO80 complex was purified from the insect cells according to a previous published protocol[16] which resulted in a pure and monodisperse sample.

INO80 complex HSA-mutants were prepared as described for WT INO80. Three Ino80(2x Flag) HSA mutants (HSAα1, HSAα2 or HSAα1/α2) were generated using standard cloning techniques and integrated into above described Baculovirus using MultiBac technology[51] (mutated residues are shown in **Supplementary Fig. 4e** and **Supplementary Table 1**).

Preparation of human mononucleosomes

Canonical human histones were essentially purified as described previously[52].

Briefly, *Escherichia Coli* BL21 (DE3) cells (Novagen) were used to express histones for 2 h at 37°C. Cells were disrupted using non-denaturating conditions and inclusion bodies were washed using 1 % Triton X-100. 7 M Guanidinium chloride was used for resuspension and inclusion bodies were dialyzed in 8 M urea. Cation exchange chromatography was applied to purify histones. After refolding of histones under low-salt conditions, an anion exchange chromatography step was used as a final purification step. Histones were lyophilized for long-time storage. To assemble histone octamers, single histones were resuspended in 7 M guanidinium chloride, mixed at a 1.2 fold excess of H2A/ H2B and dialyzed against 2 M NaCl for 16 h. Size exclusion chromatography (Superdex 200 16/600 column; GE Healthcare) was used to purify histone octamers which were then stored in 50 % glycerol at -20°C. For the purpose of mononucleosome reconstitution we used fluorescein-labeled Widom 601 DNA[53] with 80 bp extranucleosomal DNA in the 0N80 orientation[54] or without extranucleosomal DNA (0N0). After amplification by PCR, the DNA was purified using anion exchange chromatography and concentrated by applying vacuum. Histone octamers and DNA were mixed at 1.1 fold excess of DNA at 2 M NaCl. The sodium chloride concentration was then decreased to a final concentration of 50 mM over 17 h at 4°C. In a final step, NCPs were purified using anion exchange chromatography. After dialysis to 50 mM NaCl, NCPs were concentrated to 1 mg/mL and stored at 4°C.

Crystallization

Arp4-N-actin-NactNB

Prior to crystallization the Arp4-N-actin-NactNB complex (16 mg/mL) was mixed with subtilisin (1:6000 [w(protease):w(complex)] for in drop proteolysis), 0.2 mM $CaCl_2$ and either 1 mM ATP (buffered at pH 7.5 in 100 mM Tris) for the N-actin ATP bound structure or with 1 mM ADP (buffered at pH 7.5 in 100 mM Tris) for the nucleotide free (apo) structure. Crystals were grown by hanging-drop vapour diffusion at 20˚C in 1.4-1.5 M sodium malonate at pH 6.0. The best diffracting crystals were harvested after 4-8 days and cryo-protected with 23 % glycerol.

Ino80^{HSA}-Arp4-N-Actin-Arp8-module

For the crystallization of the Ino80^{HSA}-Arp4-N-Actin-Arp8-module, protein solution (13 mg/mL) was mixed with latrunculin A (LatA) (For the LatA stock solution LatA was dissolved in 100 % DMSO to a final concentration of 10 mM) at a molar ratio of 1:1.5 (Complex : LatA). Crystals were grown by hanging-drop vapour diffusion at 4˚C against 0.1 M sodium citrate tribasic dihydrate and 18 % w/v polyethylene glycol 3,350. The crystals were harvested after 30 days and cryo-protected with 20 % glycerol.

Data collection and processing, structure determination and refinement

Diffraction data from all crystals were collected at 100 K with a wavelength of 1.0 Å at the SLS (Swiss Light Source, Villigen, Switzerland) beamline X06SA. Data were processed with *XDS*[55] and scaled with *POINTLESS* and *AIMLESS* within the CCP4 suite[56].

Arp4-N-actin-NactNB

The two structures of the Arp4-N-actin-NactNB complex with N-actin ATP bound (PDB ID 5NBM) and nucleotide free (apo) (PDB ID 5NBL) were determined by molecular replacement with Phaser[57]. For a first model, structures of *S. cerevisiae* actin (PDB ID 1YAG) and Arp4 (PDB ID 3QB0) were used as search models following the removal of any nucleotides, water molecules or metal atoms. A homology model of NactNB was generated using the PHYRE server[58] and the three CDR loops were deleted prior its use as a search model. Sequential search analysis with two copies of each of the search models for Arp4, actin and NactNB resulted in a unique solution for two copies of the ternary complex per asymmetric unit. The initial model was used as search model for the analysis of the diffraction data sets from crystals grown in presence of ATP or ADP giving immediately a single solution with two complexes per asymmetric unit for both structures. In crystals grown with ATP N-actin was clearly ATP bound. In contrast in crystals grown in presence of ADP N-actin was nucleotide free. First models were then improved by iterative rounds of model refinement with *phenix.refine*[59] and manual model building with *COOT*[32]. Both electron density maps contain density for a peptide of unknown source that we could not assign to any sequence of the expressed proteins. This density was therefore modelled as a poly-UNK (unknown amino acid) peptide. The final model of the N-actin(ATP)-Arp4-NactNB complex (PDB ID 5NBM) at 3.4 Å resolution has R_{work}/R_{Free} values of 15.2/19.3 % and the model of the N-actin(apo)-Arp4-NactNB complex (PDB ID 5NBL) at 2.8 Å resolution has R_{work}/R_{Free} values of 17.1/20.4 % (**Table 1**).

Ino80^{HSA}-Arp4-N-Actin-Arp8-module

The Ino80^{HSA}-Arp4-N-Actin-Arp8-module structure (PDB ID 5NBN) was determined by molecular replacement with Phaser[57]. The Arp4-N-actin-NactNB structure (PDB ID 5NBM) without NactNB and the yeast Arp8CTD structure (PDB ID 4AM6) were used as search

models following the removal of any ligands or waters molecules. A single solution containing two copies of the Arp4-N-actin-Arp8 complex per asymmetric unit was found. Clear difference density for the Ino80^{HSA} domain was visible in the initial map after molecular replacement. The model was improved through iterative rounds of refinement with *phenix.refine*[59], applying secondary structure restraints and NCS restraints, and manual model building with COOT[32]. The Ino80^{HSA} domain was built manually with COOT[32] using B-factor sharpening and feature-enhanced-maps[60] (calculated by *phenix.fem*) for model building. Density for bound nucleotides at the canonical nucleotide binding sites of Arp4 and N-actin could be identified as ATP. Building and refinement of ADP into the unbiased density map showed in both cases clear difference density for a missing gamma-phosphate. Subsequent refinement shows similar B-factors for the alpha, beta and gamma phosphate of each ATP molecule. The final model of the Ino80^{HSA}-Arp4-N-Actin-Arp8-module model at 4.0 Å resolution has R_{work}/R_{Free} values of 19.3%/24.2% (**Table 1**).

Structures were analysed using COOT[32] and PISA[61]. Superposition of structures was performed by using the Secondary Structure Matching (SSM)[31] algorithm in COOT[32]. Figures of structures were prepared with PyMOL[62] and ChimeraX[63].

Affinity enrichment mass spectrometry (AE-MS)

Yeast with a double FLAG-tagged INO80 (Genotype: MATa INO80-FLAG$_2$ his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0)[6] were grown for 2 days in YPD medium at 30°C. Cells were harvested by centrifugation. Pellets were re-suspended 5:1 (w(yeast): w(buffer)) in 20 mM HEPES pH 7.8. The cell suspension was dripped into liquid nitrogen and the frozen cells were lysed using a freezer mill (SPEX SamplePrep). The frozen cell powder was stored at -80°C until usage.

20 g of frozen yeast cell powder was thawed in 20 mL lysis buffer (25 mM HEPES pH 8.0, 500 mM KCl, 10 % glycerol, 0.05 % NP40, 1 mM EDTA, 4 mM $MgCl_2$ and 1x PI (Sigma-Aldrich)). Chromatin was fragmented with a polytron homogenizer (Kinematica; Fisher Scientific) and by sonication (Branson). The raw cell lysate was cleared by centrifugation and 250 µg/mL avidin (IBA) was added.

The specific-binder nanobody (NactNB) and the control nanobody (enhancer GFP nanobody; eGFP-NB)[64] both had a C-terminal double Strep-Tag and were expressed and purified as described for above for NactNB. NactNB or eGFP-NB immobilized on Strep-Tactin Sepharose were incubated with equal amounts of cleared yeast cell lysate. Unbound protein was removed by washing with buffer W1 (25 mM HEPES pH 8.0, 500 mM KCl, 10 % glycerol, 0.05 % NP40, 1 mM EDTA, 4 mM $MgCl_2$) followed by buffer W2 (25 mM HEPES pH 8.0, 200 mM KCl, 10 % glycerol, 1 mM EDTA and 4 mM $MgCl_2$).

Samples for LC-MS/MS measurement were in principle prepared as published before[65]. Briefly, equal amounts of the nanobody Strep-Tactin Sepharose beads from each pull-down were incubated in buffer E1 (50 mM Tris-HCl pH 7.5, 2 M urea, 5 µg/ml trypsin (Promega) and 1 mM DTT) for 30 min at 30°C for on-bead digest. Any remaining peptides were eluted from the beads and alkylated with buffer E2 (50 mM Tris-HCl pH 7.5, 2 M urea, 5 mM iodoacetamide). Elution fractions were pooled and incubated in the dark overnight at 32°C. The digestion was stopped by the addition of 1 % trifluoroacetic acid. Samples were loaded on self-made C18 reversed-phase StageTips for purification and enrichment following a standard protocol[66]. Peptides were eluted with 2x20 µL buffer B (80 % ACN and 0.5 % AcOH) and concentrated using a SpeedVac concentrator to a final volume of 5-10 µL. Finally, 2.5 µL of buffer A* (2 % ACN, and 1 % TFA) and 2.5 µL buffer A (0.5 % AcOH) were added to the sample.

Peptide samples were measured on an LC-MS/MS system using a UHPLC (EASY-nLC 1000) coupled to an LTQ Orbitrap Elite (both Thermo Scientific) equipped with a standard nanoelectrospray source. Peptides were loaded onto a 15 cm × 0.050 mm I.D. reversed phase column packed with 2 µm C18 beads (Acclaim PepMap RSLC analytical column, Thermo Scientific) and subsequently separated using a 90 min gradient of solvent B (98 % ACN, 0.1 % FA) from 2 % to 35 % at a flow rate of 250 nl/min.

*.RAW files from the eGFP-NB (mock) and NactNB triplicate experiments were analysed together using the MaxQuant software suite (version 1.5.2.18) including the label-free algorithm for LFQ intensity calculation[67]. Downstream data analysis was performed in the Perseus environment (version 1.5.0.9.)[68]. Briefly, LFQ intensity values were $\log_{10}$ transformed, the data were filtered for at least 2 valid values in at least one of the two conditions and missing values were imputed using a normal distribution at the noise level (width: 0.3 of the standard deviation of the data; down shift: 1.8 standard deviations of the valid data). To reveal significant outliers, a two-sample t-test was performed and data were visualized using an in-house R script.

Fluorescence anisotropy

Arp8-module in solution 40 bp dsDNA binding affinity was measured by fluorescence anisotropy in principle as described before[69].

Equimolar amounts of the two complementary DNA strands (forward 5`-3`: fluorescein-CCCTGGCGACTTCGCCTCGTTTTGGCGATTTTCTTAGCAAATATTCTTTC and reverse 5`-3`: GAAAGAATATTTGCTAAGAAAATCGCCAAAACGAGGCGAAGTCGCCAGGG), solved in water, were heated to 95°C for 10 min and slowly cooled at room temperature to anneal the two DNA strands. Arp8-module was diluted to the respective working concentration and incubated with 20 nM dsDNA on ice for 30 min in 20 mM Tris pH 7.8, 50 mM KCl and 2.5 %

glycerol in a total volume of 100 µL. Fluorescence anisotropy was measured in a black flat-bottomed non-binding 96 well plate (Greiner-Bio) on a Tecan Infinite M1000 plate reader (Excitation wavelength 470 nM; Emission wavelength 520 nM).
Data was analysed and fitted to a non-linear non-cooperative 1:1 binding model (y = Af - (Af-Ab) * ((x) / (Kd + x); y anisotropy; Af anisotropy of free ligand; Ab Anisotropy of bound ligand; Kd dissociation constant; x receptor concentration) with the program Prism (GraphPad) to calculate the dissociation constants for the respective complex. Experiments were performed in triplicates.

Electro mobility shift assays

The Arp8-module binding preference for mononucleosomes with or without extranucleosomal DNA was examined with competition electro mobility shift assays.
Increasing amount of Arp8-module were titrated against a 1:1 mixture of 0N0 and 0N80 (20 nM each) mononucleosomes in 10 mM HEPES pH 8.0, 2 mM $MgCl_2$, 60 mM NaCl, 8 % glycerol and incubated for 20 min on ice. 15 µL of each titration step were loaded on a precast native polyacrylamide gel (NativePAGE Novex™ 4-16 % Bis-Tris Protein Gels; Invitrogen). Arp8-module bound and unbound nucleosomes were resolved by Native-PAGE in 1x NativePAGE™ Running Buffer (Invitrogen; according to the manufacturer protocol) at 120 V for 120 min at 4°C. Gels were analysed on a Typhoon FLA 9000 plate reader (GE Healthcare) with 25 µm pixel size, using FITC fluorescence scan.
To test the binding capability of INO80 to nucleosomes a titration of the complex was carried out. Increasing amounts of the protein in 25 mM Hepes, pH 8.0, 60 mM KCl, 7 % glycerol, 1 mM $CaCl_2$ were incubated with 20 nM 0N80 nucleosomes for 30 min on ice. INO80 bound and unbound nucleosomes were resolved by NativePAGE (Novex™ 4-16 % Bis-Tris Protein Gels; Invitrogen) and subsequently visualized on a Typhoon FLA 9000 plate reader as described above.

Nucleosome sliding assays

The nucleosome sliding activity of INO80 was monitored on 0N80 mononucleosomes.
18 nM INO80 was incubated with 90 nM 0N80 nucleosome in sliding buffer (25 mM Hepes, pH 8.0, 60 mM KCl, 7 % glycerol, 0.10 mg/mL BSA, 0.25 mM DTT, 2 mM $MgCl_2$) at 26°C. The sliding reaction was started by the addition of ATP and $MgCl_2$ (final concentrations: 1 mM ATP and 2 mM $MgCl_2$). At the respective time points (30, 60, 120, 300, 600, 1800 and 3600 s) the reaction was stopped by adding lambda DNA (NEB) to a final concentration of 0.2 mg/mL. Native PAGE (NativePAGE™ Novex™ 4-16% Bis-Tris Protein Gels; Invitrogen) was used to separate distinct nucleosome species. Gels were visualized on a Typhoon FLA 9000 plate reader as described above.

ATPase assays

In order to determine the ATPase rate of INO80 we applied an NADH-based ATPase assay in principle as described in [24,70].

Briefly, 27 nM of INO80 was incubated in assay buffer (25 mM Hepes, pH 8.0, 50 mM KCl, 5 mM $MgCl_2$, 0.1 mg/mL BSA) with 0.5 mM phosphoenolpyruvate, 2 mM ATP, 0.2 mM NADH and 25 U/mL lactate dehydrogenase/pyruvate kinase (Sigma) in a final volume of 50 µL at 30°C. The Tecan Infinite M100 (Tecan) was used to monitor the NADH dependent fluorescence signal in non-binding, black 384-well plates (Greiner) at an excitation wavelength of 340 nm and an emission wavelength of 460 nm over a time course of 40 min. ATPase activity for all samples was determined at conditions of maximum INO80 WT ATPase activity. Stimulation was performed with 50 nM 0N80 nucleosome, 100 nM 0N0 nucleosome or 100 nM 223 bp DNA (DNA template used to reconstitute 0N80 nucleosomes). The final ATP turnover rate was calculated using maximal initial linear rates which were corrected for a buffer blank.

The **Genome-wide in vitro reconstitution assay** and the **Restriction enzyme accessibility assay** are described in the **Supplementary Notes**.

Data availability statement

Coordinates and structure factors have been deposited in the Protein Data Bank under PDB ID accession codes **5NBM** for the N-actin(ATP)-Arp4-NactNB, **5NBL** for the N-actin(apo)-Arp4-NactNB and **5NBN** for the $Ino80^{HSA}$-Arp4-N-Actin-Arp8-module structures.

Data of the genome-wide nucleosome positioning experiments reported in this paper have been deposited on the NCBI Gene Expression Omnibus (accession number **GSE113401**).

All other data and material are available from the corresponding author upon reasonable request.

Methods references

49. Rothbauer, U. et al. Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat Methods* **3**, 887-9 (2006).
50. Conrath, K.E. et al. Beta-lactamase inhibitors derived from single-domain antibody fragments elicited in the camelidae. *Antimicrob Agents Chemother* **45**, 2807-12 (2001).
51. Trowitzsch, S., Bieniossek, C., Nie, Y., Garzoni, F. & Berger, I. New baculovirus expression tools for recombinant protein complex production. *J Struct Biol* **172**, 45-54 (2010).
52. Dyer, P.N. et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol* **375**, 23-44 (2004).
53. 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).
54. Levendosky, R.F., Sabantsev, A., Deindl, S. & Bowman, G.D. The Chd1 chromatin remodeler shifts hexasomes unidirectionally. *Elife* **5**(2016).
55. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32 (2010).
56. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-42 (2011).
57. McCoy, A.J. et al. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).
58. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. & Sternberg, M.J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-58 (2015).
59. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).
60. Afonine, P.V. et al. FEM: feature-enhanced map. *Acta Crystallogr D Biol Crystallogr* **71**, 646-66 (2015).
61. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97 (2007).
62. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).
63. Goddard, T.D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018).
64. Kirchhofer, A. et al. Modulation of protein properties in living cells using nanobodies. *Nat Struct Mol Biol* **17**, 133-8 (2010).
65. Keilhauer, E.C., Hein, M.Y. & Mann, M. Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol Cell Proteomics* **14**, 120-35 (2015).
66. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896-906 (2007).
67. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-26 (2014).
68. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-40 (2016).

69. Favicchio, R., Dragan, A.I., Kneale, G.G. & Read, C.M. Fluorescence spectroscopy and anisotropy in the analysis of DNA-protein interactions. *Methods Mol Biol* **543**, 589-611 (2009).
70. Kiianitsa, K., Solinger, J.A. & Heyer, W.-D. NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Analytical Biochemistry* **321**, 266-271 (2003).

Appendix

The following chapters are not published in a peer-reviewed journal yet, but they are under revision in Nature Communications.

CHAPTER 3: GENOME INFORMATION PROCESSING BY THE INO80 CHROMATIN REMODELER POSITIONS NUCLEOSOMES

Elisa Oberbeckmann[1,2,8], Nils Krietenstein[1,3,8], Vanessa Niebauer[4], Yingfei Wang[5], Kevin Schall[4], Manuela Moldt[4], Tobias Straub[6], Remo Rohs[5], Karl-Peter Hopfner[4], Philipp Korber[1], Sebastian Eustermann[4,7]

[1]Division of Molecular Biology, Biomedical Center, Faculty of Medicine, Ludwig-Maxmilians-Universität München, Martinsried near to Munich, Germany; [2]current address: Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany; [3]current address: Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts, USA; [4]Gene Center, Faculty of Chemistry and Pharmacy, Ludwig-Maxmilians-Universität München, Munich, Germany; [5]Quantitative and Computational Biology, Departments of Biological Sciences, Chemistry, Physics & Astronomy, and Computer Science, University of Southern California, Los Angeles, CA, USA; [6]Core Facility Bioinformatics, Biomedical Center, Faculty of Medicine, Ludwig-Maxmilians-Universität München, Martinsried near to Munich, Germany; [7]current address: European Molecular Biology Laboratory (EMBL), Structural and Computational Biology Unit, Heidelberg, Germany; [8]These authors contributed equally.

This chapter is not published in a peer-reviewed journal yet
but under revision and made public on BioRxiv.

Author contributions to "Genome information processing by the INO80 chromatin remodeler positions nucleosomes"

Conceptualization: E.O., N.K., T.S., R.R., K.-P.H., P.K., S.E.; Data curation: E.O., N.K., V.N.; Formal analysis: E.O., N.K., V.N., Y.W., T.S.; Funding acquisition, Project administration, Supervision: K.-P.H., P.K., S.E.; Investigation: E.O., N.K., V.N., K.S., M.M., Y.W., T.S., S.E.; Methodology: E.O., N.K., V.N., K.S., R.R., T.S., P.K., S.E.; Validation: E.O., N.K., V.N., K.S., Y.W., R.R., T.S., P.K., S.E.; Visualization: E.O., N.K., V.N., Y.W., T.S., P.K., S.E.; Writing original draft: E.O., N.K., P.K., S.E.; Writing – review & editing: E.O., N.K., V.N., T.S., R.R., K.-P.H., P.K., S.E.

---

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Genome information processing by the INO80 chromatin remodeler positions nucleosomes

Elisa Oberbeckmann[1,2,8], Nils Krietenstein[1,3,8], Vanessa Niebauer[4], Yingfei Wang[5], Kevin Schall[4], Manuela Moldt[4], Tobias Straub[6], Remo Rohs[5], Karl-Peter Hopfner[4]*, Philipp Korber[1]*, Sebastian Eustermann[4,7]*

[1]Division of Molecular Biology, Biomedical Center, Faculty of Medicine, Ludwig-Maxmilians-Universität München, Martinsried near to Munich, Germany; [2]current address: Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany; [3]current address: Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts, USA; [4]Gene Center, Faculty of Chemistry and Pharmacy, Ludwig-Maxmilians-Universität München, Munich, Germany; [5]Quantitative and Computational Biology, Departments of Biological Sciences, Chemistry, Physics & Astronomy, and Computer Science, University of Southern California, Los Angeles, CA, USA; [6]Core Facility Bioinformatics, Biomedical Center, Faculty of Medicine, Ludwig-Maxmilians-Universität München, Martinsried near to Munich, Germany; [7]current address: European Molecular Biology Laboratory (EMBL), Structural and Computational Biology Unit, Heidelberg, Germany.

The fundamental molecular determinants by which ATP-dependent chromatin remodelers organize nucleosomes across eukaryotic genomes remain largely elusive. Here, chromatin reconstitutions on physiological, whole-genome templates reveal how remodelers read and translate genomic information into nucleosome positions. Using the yeast genome and the multi-subunit INO80 remodeler as a paradigm, we identify DNA shape/mechanics encoded signature motifs as sufficient for nucleosome positioning and distinct from known DNA sequence preferences of histones. INO80 processes such information through an allosteric interplay between its core- and Arp8-modules that probes mechanical properties of nucleosomal and linker DNA. At promoters, INO80 integrates this readout of DNA shape/mechanics with a readout of co-evolved sequence motifs via interaction with general regulatory factors bound to these motifs. Our findings establish a molecular mechanism for robust and yet adjustable +1 nucleosome positioning and, more generally, remodelers as information processing hubs that enable active organization and allosteric regulation of the first level of chromatin.

The packaging of DNA with histones into nucleosomes underpins the maintenance and regulation of genome information in eukaryotes[1,2]. Genome-wide mapping of chromatin revealed highly-defined patterns of nucleosomes carrying a combinatorial landscape of histone variants and modifications[3-8]. These patterns entail well-positioned nucleosomes, which occupy the same genomic position across a cell population and even adopt equivalent positions relative to genomic sites of equivalent function like transcription start sites (TSS)[6,7]. Most prominently, nucleosome-depleted regions (NDRs) at promoters of active or poised genes are flanked by a well-positioned hallmark nucleosome (+1 nucleosome) that is the first in a regular nucleosome array over the transcribed region[9]. These stereotypic NDR-array patterns are conserved from yeast to man, and changes within their configuration play a pivotal role in transcriptional regulation, e.g., during cell differentiation and stress response[10,11]. Understanding the fundamental molecular determinants of nucleosome positioning is likely to reveal core principles by which genome regulation occurs.

A nucleosome position is defined by the DNA sequence that is wrapped around the histone octamer[12]. While this DNA sequence always answers the question “Where is this nucleosome?”, it may, but need not, answer the question “How was the nucleosome placed there?”. Histone octamers may form nucleosomes virtually at any DNA sequence position in the genome[13]. A molecular mechanism that consistently places a nucleosome at a particular genome position across a cell population must select this position against competing positions. This selection may be based on genetic information encoded within DNA sequence or on epigenetic information like histone modifications and variants or other chromatin-associated factors. Regarding DNA sequence information, pioneering studies proposed two mechanisms (Fig. 1a). One

[8]These authors contributed equally.

* Corresponding authors: sebastian.eustermann@embl.de, pkorber@lmu.de, hopfner@genzentrum.lmu.de

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mechanism relies on the intrinsic specificity of nucleosomes to preferentially assemble on DNA sequences that favor the wrapping around the histone octamer ("genomic code for nucleosome positioning")[14,15]. In this case, the nucleosomal DNA sequence directly determines the position. The other mechanism requires DNA sequence-specific binding of a barrier factor, to which one or several nucleosomes are aligned by statistical positioning regardless of the octamer-bound DNA sequences[16]. The principal difference between these two mechanisms illustrates two extremes, which pertain to the central question whether DNA sequence information directly or indirectly determines a nucleosome position. If directly, the nucleosome positioning mechanism reads out the DNA sequence information at the resulting nucleosome position itself. If indirectly, DNA sequence is read somewhere else, and the resulting positioning information is relayed by alignment mechanisms that position nucleosomes relative to barriers and other nucleosomes. In this case, the DNA sequence bound by the histone octamer would define, but not directly determine, the genomic position of a nucleosome.

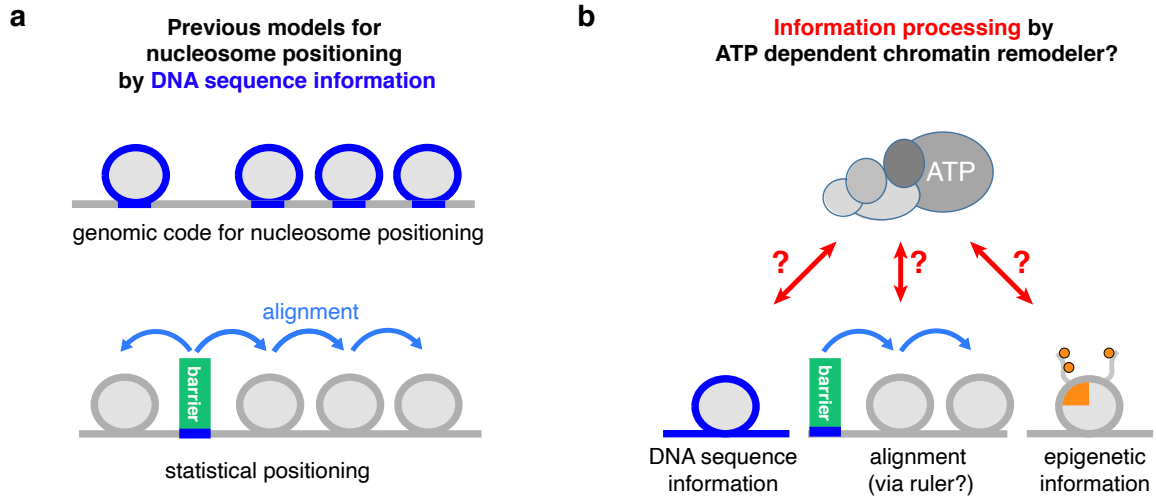


Figure 1. Models for nucleosome positioning mechanisms. a Genomic code for nucleosome positioning[14,15] and statistical positioning[16] are two previous models, which exemplify a direct versus indirect *role, respectively,* of DNA sequence information for determining nucleosome positioning. **b** In light of the decisive role of ATP-dependent chromatin remodelers in nucleosome positioning[24,28,29,67], we asked if and how these large, macro-molecular machines actively process (epi)genetic information together with their own remodeler-specific information into stable nucleosome positioning.

In recent years, it has become clear that the pure versions of these two mechanistic extremes fail to explain nucleosome positioning *in vivo*. Intrinsic histone octamer preferences, as operationally assessed by salt gradient dialysis (SGD) reconstitution from purified DNA and histones[13], cannot recapitulate NDR-array patterns *in vitro*[17,18], and inter-nucleosomal distances (spacing) are independent of nucleosome density *in vivo*[19,20] and *in vitro*[18,21] in contrast to predictions of the statistical positioning mechanism[16,22].

Instead, ATP-dependent chromatin remodelers have now been established as decisive nucleosome positioning factors by studies both *in vivo* and *in vitro*. Chromatin remodelers often form multisubunit macromolecular complexes and are grouped into four families: INO80/SWR1, SWI/SNF, ISWI, CHD. By using energy derived from ATP hydrolysis, remodelers alter histone-DNA interactions resulting in nucleosome translocation (sliding), ejection, and reconfiguration[23]. Mutations in genes encoding remodeler subunits, especially combined mutations, lead either to compromised nucleosome patterns and composition, or are lethal[20,24-28]. Complementary to genetic studies, cell-free reconstitutions provided direct evidence for the critical role of chromatin remodelers in nucleosome positioning and allowed to distinguish remodeler contributions from those of other factors, like the transcription and replication machinery[18,29]. Nucleosomes were assembled by SGD, even for an entire genome with yeast genomic DNA fragments or plasmid libraries[17,18,29,30]. The largely non-physiological nucleosome positions generated by SGD were turned in an ATP-dependent manner into *in vivo*-like NDR-array patterns either by addition of whole cell extracts[18] or, remarkably, also by addition of remodelers purified from yeast[29]. For example, addition of yeast INO80 or SWI/SNF-type RSC remodeling complexes to SGD chromatin generated hallmark features of *in vivo*-like nucleosome organization, +1 nucleosomes and NDRs at

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

promoters, respectively[29]. This argued for a remodeler-mediated direct readout of positioning information, possibly involving DNA sequence features[29,31] and epigenetic information[23]. Notably, various remodelers contain reader domains of histone marks, while most of them lack classical sequence-specific DNA binding domains. This led to the proposal that remodelers, similar to histones, may recognize sequence dependent structural features of DNA such as DNA shape[32,29]. Ample and growing evidence for transcription factors underscores the functional relevance of DNA shape features in genome regulation[33]. Such features might be relevant at poly(dA:dT)-rich promoter sequences, which have been implicated in regulation of RSC activity at the NDR[31,29], while we hypothesized that DNA shape might also play a role during +1 nucleosome positioning by INO80[29]. In contrast, other remodelers, such as the yeast ISW1a and ISW2 complexes could not generate *in vivo*-like nucleosome positions on their own but required sequence readout by other factors. So-called "general regulatory factors" (GRFs) are sequence-specific DNA binding proteins, often essential for viability and involved in transcription or replication regulation via their impact on chromatin organization[34-36]. Addition of purified GRFs, e.g., yeast Reb1 or Abf1, enabled the ISW1a and ISW2 remodelers to align regular nucleosome arrays relative to the GRF binding sites[29]. This argued in turn for remodeler-mediated readout of sequence information via processive alignment at GRFs as well as among nucleosomes, possibly involving a protein ruler[37].

Although cell-free reconstitution and genetic studies established the critical importance of remodelers in determining the genomic organization of nucleosomes, the dissection of the underlying molecular mechanism and the required information has proven difficult. Recent structural work shed light onto the architecture of different remodelers and how they might act on mono-nucleosomes [38]. However, there remains the conundrum that the principal remodeler activity of mobilizing nucleosomes must be regulated such that it results in stable nucleosome positions relative to genomic sequence.

In this study, we directly addressed this fundamental conundrum by asking which kind of DNA sequence, histone, barrier or other epigenetic information provides the required input, and how remodelers turn this information input into stable nucleosome positioning (Fig. 1b). We advanced whole genome reconstitutions into a fully recombinant, *de novo* approach. In this system full biochemical control is established by using recombinant components in conjunction with high resolution structural information enabling the identification of remodeling mechanisms. Not only the core mechanism of remodelers, as studied so far mainly in mono-nucleosome assays, but also the extended functions arising from remodeling of chromosomal multi-nucleosome substrates as well as the readout of physiological genomic DNA sequences and other nucleosome positioning information can be assessed at a detailed mechanistic level. We used the yeast genome and the multi-subunit structure of the INO80 complex as a paradigm to identify and probe the information and mechanism by which remodelers read information and translate it into stable nucleosome positions. In the accompanying study (Oberbeckmann & Niebauer et al.), we addressed how remodelers propagate nucleosome positioning information via an alignment mechanism to generate phased and regular nucleosomal arrays. Taken together, our data reveal that and how remodelers are information processing hubs. Genome information encoded within DNA shape/mechanics as well as in DNA sequence motifs bound by barrier factors is actively read out by the remodelers and integrated via the allosteric interplay of their molecular machinery into nucleosome positions.

Results

A fully recombinant approach for *de novo* whole-genome reconstitutions. To explore how ATP-dependent chromatin remodelers place nucleosomes at *in vivo*-like positions, we advanced whole-genome reconstitutions[18,29,30] into a fully recombinant *de novo* approach (Fig. 2a). We established recombinant production of highly active and stoichiometric INO80 complex (Supplementary Fig. 1a,b) and performed whole-genome reconstitutions using recombinant histones and a fully-defined yeast genomic plasmid library[39]. This leverages, compared to previously used ill-defined plasmid libraries, endogenous fly embryo histones and endogenous purifications of remodelers[29], the full potential of biochemical systems: (1) A fully defined 15-subunit *S. cerevisae* INO80 complex, amendable for structure-guided mutagenesis, (2) histones without posttranslational modifications (PTMs) and amendable for mutagenesis, and (3) defined DNA templates for chromatin assembly. We used MNase-seq to measure resulting nucleosome positions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

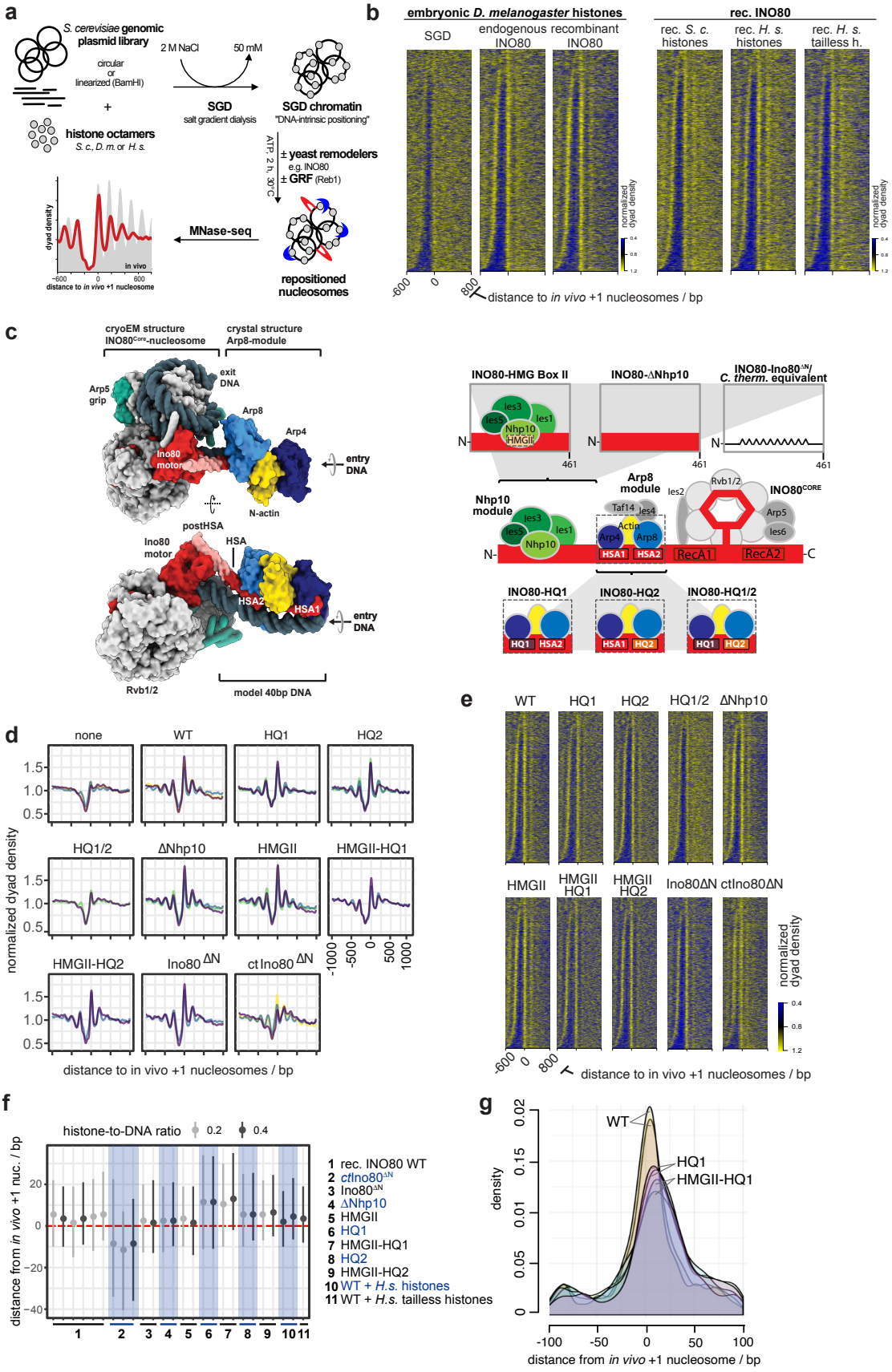


Figure 2. Fully recombinant genome-wide reconstitution of nucleosome positioning by INO80. a Overview of genome-wide *in vitro* chromatin reconstitution system. **b** Heat maps of MNase-seq data for SGD chromatin assembled with embryonic or recombinant (rec.) histones from the indicated species ("*H. s.*" abbreviates *Homo sapiens*, "*S. c.*" abbreviates *Saccharomyces cerevisiae*.) and remodeled with endogenous or recombinant *S. cerevisiae* INO80 complex as indicated. Heat maps are aligned at *in vivo* +1 nucleosome positions and sorted by NDR length. Single replicates are plotted, see Supplementary Figure 1c for

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

all replicates. **c** Left panel: Composite model of INO80 based on high resolution cryoEM structure of ctINO80 core in complex with a mono-nucleosome[43] and X-ray structure of Arp8 module modeled on 40bp linker DNA[46]. Images taken from Knoll et al. [46]. Right panel: Schematic of INO80 complex submodule and subunit organization (middle) with zoom into Nhp10 (top) or Arp8 module (bottom) showing three mutant versions each. **d** Composite plots of MNase-seq data of individual replicates for SGD chromatin incubated with the indicated recombinant WT (WT) or mutant INO80 complexes (as in panel c) from *S. cerevisiae* or *C. thermophilum* (*ct*INO80$^{\Delta N}$). **e** Heat maps of MNase-seq data for samples as in panel d. **f** Distributions of distances between +1 nucleosome positions determined by paired-end sequencing after reconstitution by the indicated combinations of INO80 complexes and histones at the indicated histone-to-DNA mass ratio relative to *in vivo* +1 nucleosome positions. Dots mark the medians, vertical lines the interquartile distances. Alternating white and grey vertical zones group replicates of the indicated remodeler/histone combinations. **g** Density distributions of MNase-seq reads relative to *in vivo* +1 nucleosome positions of samples with INO80 WT, HQ1 and HMGII-HQ1 mutant complexes as in panel f.

DNA sequence and globular histone octamer information is sufficient for *in vivo*-like +1 nucleosome positioning by INO80. This recombinant system enabled us to identify the minimal information for nucleosome positioning by INO80. Consistent with its localization and function *in vivo*[40], INO80 positions *in vivo*-like +1 nucleosomes adjacent to NDRs (Fig. 2b,[29]). As equally pronounced +1 nucleosome positioning activity was observed for recombinant as for endogenous INO80 (Fig. 2b, left), we concluded that no yeast-specific PTMs were required and no co-purified yeast contaminant was responsible. To control the specificity of the highly pure INO80 complex (Supplementary Fig. 1a,b), we assayed an INO80 complex which carries a Walker B motif mutation within its Ino80 ATPase motor protein (Supplementary Fig. 1c) and excluded that nucleosome positioning activity was due to any co-purifying factor(s) from insect cells. Intriguingly, our recombinant whole-genome reconstitutions established conditions, under which INO80 generated extensive nucleosome arrays (e.g., upon addition of Reb1, see below). This served as starting point for the study of nucleosome spacing mechanisms (accompanying paper by Oberbeckmann & Niebauer et al.).

Next, we asked whether epigenetic information derived from histone modifications or variants was required for +1 nucleosome positioning. Histone variants, for example H2A.Z, may alter direct, sequence-dependent interactions of the histone octamer[41]. However, compared to SGD chromatin prepared with endogenous fly histones, using either recombinant human or yeast histones resulted in very similar nucleosome positioning by INO80 (Fig. 2b, right). Patterns were less pronounced with yeast histones, which we attributed to their known propensity to form less-stable nucleosomes[42]. As the species-origin of the histones did not matter much, we went more minimalistic and asked if just the globular histone domains were sufficient. SGD chromatin with recombinant tailless human histones still allowed INO80 to position *in vivo*-like +1 nucleosome position (Fig. 2b, right). We observed increased sliding rates with tailless compared to full-length histone nucleosomes (Supplementary Fig. 1d) consistent with previous studies[43-45]. Nonetheless, this increased sliding rate did not abrogate formation of the steady state nucleosome positioning pattern.

Taken together, we concluded that neither histone modifications nor histone variants nor histone tails nor yeast-specific modifications are absolutely required for INO80 principal activity to position in vivo-like +1 nucleosome. Consequently, INO80 can generate such positioning solely by processing information from genomic DNA sequences and the globular histone octamer. Nonetheless, a readout of epigenetic information by remodelers is expected to play a pivotal role in the regulation of nucleosome positioning, e.g., in response to changes in the cellular environment, as discussed further below.

Structure-based site-directed mutagenesis probes nucleosome positioning by INO80. Having identified a minimal set of components, from which INO80 derives nucleosome positioning information, we set out to specify this information and to dissect the molecular mechanism, by which it was processed. To this end, we leveraged high-resolution structures of INO80[43,45,46] and asked which remodeler elements might function as reader of genome information.

Recent structural and biochemical studies revealed an extended configuration of the INO80 multi-subunit architecture on mono-nucleosomes (Supplementary Fig. 1f): the INO80 core module (Ino80 protein containing the Snf2-type ATPase, Ies2, Ies6, Arp5, Rvb1, Rvb2) engages the nucleosome core particle[43,45], the nuclear-actin containing Arp8 module (Ino80-HSA domain, Arp8, Arp4, nuclear actin, Ies4 and Taf14) binds along 40-50 bp of linker DNA at the entry site[43,45,47], while the species-specific Nhp10 module (Nhp10, Ies1-3 and Ies5) bound to the Ino80 N-terminal region is located at the distal site of INO80's linker DNA footprint[47]. Linker DNA binding by the Arp8 and Nhp10 modules was proposed to provide a DNA linker length dependent sensor that is allosterically coupled to processive nucleosome translocation

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

catalyzed by the INO80 core[46-48]. *In vivo* ChIP-exo mapping suggested a highly similar INO80 configuration at +1 nucleosomes with the Arp8 or Nhp10 modules located at adjacent promoter regions[40]. Thus, we reasoned that these INO80 modules are prime candidates for reading genomic DNA sequence information.

To test this hypothesis, we targeted candidate INO80-DNA interactions based on the high-resolution cryoEM and X-ray structures of the INO80 core and Arp8 module, respectively, as well as on homology modeling of the structurally less well characterized Nhp10 module. For the INO80 core, we tested the role of ATP hydrolysis by the hetero-hexameric AAA$^+$-ATPase Rvb1/2 (Fig. 2c, Supplementary Fig. 1c), which structurally organizes the nucleosome core binding and remodeling unit of INO80[43,45]. For the Arp8-module, we employed the Ino80-HSA helix mutants, which contain substitutions of highly conserved lysine/arginine to glutamine residues in the HSAα1 and/or HSAα2 helices (HQ1, HQ2 and combined HQ1/2 mutants, respectively) that are important for linker DNA binding[46] (Fig. 2c, Supplementary Fig. 1e). For the Nhp10 module, we either mutated site-specifically the HMG box II in Nhp10 based on well-known DNA binding activity of HMG box proteins or removed the entire Nhp10 module by deleting Nhp10 or truncating Ino80's N-terminal 1-461 residues, to which this module binds (Fig. 2c, Supplementary Fig. 1e,g,h). This latter mutant corresponded to the *Chaetomium thermophilium* INO80 core complex used in the cryoEM structure[43], which we also employed here. Nhp10 module HMGII box and Arp8-module HQ1 or HQ2 mutations were also combined (HMGII-HQ1, HMGII-HQ2 mutants, respectively) (Fig. 2c, Supplementary Fig. 1e).

The INO80 Arp8 module is a reader of genomic sequence information. Comparison of nucleosome patterns in aligned heat map or composite plots suggested that most INO80 mutant complexes generated similar +1 nucleosome positioning as WT INO80 (Fig. 2d,e, Supplementary Fig. 1c). Rvb1/2 ATPase activity was not required (Supplementary Fig. 1c), consistent with the likely role of Rvb1/2 during INO80 biogenesis[49]. Even the heterologous *C. thermophilum* INO80 core complex (ctINO80$^{\Delta N}$) appeared to generate +1 nucleosomes on the *S. cerevisiae* genome to a remarkable extent, suggesting a conserved readout mechanism (Fig 2d,e). Only the HQ1/2 double mutant complex was substantially impaired in +1 nucleosome positioning (Fig. 2d,e), consistent with its impaired nucleosome sliding and decoupled ATPase activity[46]. The apparent robustness of INO80 +1 nucleosome positioning activity was in contrast to the nucleosome spacing activity, which was affected for most of these INO80 mutants (accompanying paper by Oberbeckmann & Niebauer et al.).

Quantification of distances between +1 nucleosome positions reconstituted *in vitro* and observed *in vivo* revealed a distinct impact of INO80 mutations (Fig. 2f,g). Paired-end sequencing enabled accurate determination of nucleosome dyad positions on individual DNA molecules, while we included also a lower histone-to-DNA mass ratio (~0.2, accompanying paper by Oberbeckmann & Niebauer et al.) than mostly used in this study (~0.4) to further reduce possible next-neighbor nucleosome effects. WT INO80 and Nhp10 module mutants generated *in vivo*-like +1 nucleosomes with remarkable precision (Fig. 2f,g), whereas INO80 complexes bearing the HQ1 mutation and the ctINO80$^{\Delta N}$ complex generated +1 nucleosome positions that deviated more from the *in vivo* positions than those generated by the other complexes (Fig. 2f). Compared to WT INO80, +1 nucleosome positioning by complexes with the HQ1 mutation was shifted by 10 bp downstream and reduced positioning precision was reflected in broadened distributions, which suggests that DNA sequences underlying *in vivo* +1 nucleosome positions correspond more to the DNA sequence preferences for nucleosome positioning of the WT versus the mutant INO80 complexes (see below). (Fig. 2f,g). Such downstream shifts, observed here for individual INO80 point mutations, were reminiscent of similar effects resulting from INO80 depletion in the context of the interplay with other remodelers *in vivo*[20,28,40,50].

Taken together, our mutational analysis of candidate DNA contacts indicated robust processing of genomic sequence information by INO80 with a decisive role of the Arp8, but not the Nhp10 module, as direct reader of genome information at promoters.

DNA shape/mechanics readout underlies nucleosome positioning by INO80. Based on our mutational analysis, we sought to identify genomic DNA sequence features that provide positioning information. Previously, we proposed that *S. cerevisiae* INO80 might read DNA shape features of nucleosomal DNA[29]. However, this hypothesis was based on correlation and the approach limited further interpretation, mainly because we used gene ranking by MNase-seq signal strength at pre-defined +1 to +3 nucleosome regions before and after remodeling as the discriminating category. This may introduce a bias towards the starting conditions, i.e. DNA sequence preferences of histones and variations in SGD assembly conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Moreover, the analysis was limited to pre-defined regions and numerous other DNA sequence motifs present at gene starts, e.g., evolved in the context of transcription regulation, may have convoluted the search for positioning information.

Here, we overcame these limitations and searched for the DNA sequence features of nucleosome positioning preferences by INO80 more globally, not only at promoters, and explored by a structure-based mutational analysis the direct and causal impact of altered INO80-DNA contacts on these preferences. We established a sensitive and unbiased Principal Component Analysis (PCA)/clustering approach solely on the basis of *de novo* generated nucleosome dyad positions determined by paired-end sequencing. This enabled unsupervised PCA/clustering of a large number of datasets (e.g. replicates, different assembly degrees, various INO80 WT and mutant complexes etc.) without prior assumptions (Fig. 3a).

Nucleosomes remodeled by WT INO80 clearly clustered differently in PCA than those assembled during SGD without remodeling (Fig. 3b), i.e. this approach could clearly distinguish positioning preferences under different conditions. The DNA sequences of different clusters did not differ in terms of sequence motifs assessed by motif search algorithms like Homer (data not shown) in contrast to previous studies of an isolated, truncated construct of human INO80 HSA domain that indicated sequence-specific DNA binding[51].

However, DNA sequence information need not result in classical sequence motifs but may correspond to DNA shape features that are encoded in a more redundant way, i.e., rather disparate sequences may share similar shape features[52]. A composite plot of the DNA shape feature propeller twist of SGD-reconstituted versus INO80-remodeled nucleosomes revealed symmetrical but strikingly different profiles (Fig. 4a), revealing distinct DNA sequence requirements for INO80- and SGD-mediated positioning. Whereas propeller twist is largely affected by the number of intra-bp hydrogen bonds, other shape features gave corresponding results (Supplementary Fig. 2a). These other shape features take into account interactions either between adjacent bp (helix twist and roll) or with additional nucleotides (minor groove width). The profile symmetry validated the shape information content as no nucleosome orientation was to be expected and symmetrical shape profiles are unlikely to occur by chance if no underlying shape feature were involved. Importantly, similar but asymmetrically distorted shape profiles were seen for nucleosomes reconstituted at positions close to *in vivo* +1 nucleosome positions and oriented relative to the direction of transcription (Fig. 4c). This shows that such pronounced DNA shape signals are also present in +1 nucleosome regions at gene promoters and strongly suggested that we identified the DNA-encoded signal for INO80-mediated +1 nucleosome positioning. The structural readout of DNA features, both in the gene promoter as well as in +1 nucleosome, is also consistent with *in vivo* binding of INO80 subunits to such regions, as observed by ChIPexo mapping[40].

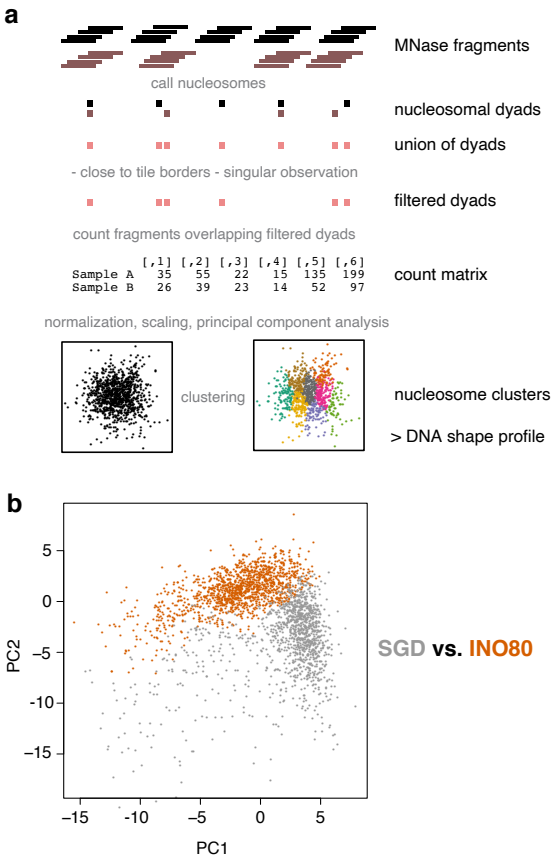


Figure 3. Principal Component Analysis (PCA)/clustering approach. a Schematic of the analysis by using two conditions (black and grey) as an example. For details see main text and Materials and Methods section. **b** Visualization of nucleosome clusters according to Principal Components 1 and 2 (PC1, PC2) for SGD chromatin (SGD) prepared with embryonic *D. melanogaster* histones at histone-to-DNA mass ratio of 0.4 alone (SGD) or after incubation with *S. cerevisiae* WT INO80 complex (INO80). INO80 remodeling alters almost the entire landscape of nucleosome positions.

DNA shape profiles establish a new kind of nucleosome positioning information that is distinct from previously known DNA sequence preferences of histones. The relevance of DNA shape for remodeler-mediated nucleosome positioning was further underscored by a striking congruency between our PCA/clustering data and high-resolution structural information as well as *in vivo*-mapping of INO80 subunits at gene promoters. The

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

remodeled nucleosomes differed mostly in the ± 55 bp and ± 100 bp regions relative to the dyad (color shaded areas in Fig. 4a) where functionally important interactions with the INO80 complex are suggested by the biochemical and structural information available from INO80 in complex with mono-nucleosomes (Fig. 4b). The HSA helix at the Ino80 N-terminus contacts linker DNA at about -100 bp from the dyad[46,47]. The -55 bp region from the dyad lies between the Ino80 ATPase domain and the DNA contact point of Arp5. Both of these regions are critically important for nucleosome translocation. DNA strain build-up in the -55 bp region by successive rounds of DNA pumping by Ino80 ATPase motor is a central element of the proposed core mechanism of nucleosome translocation by INO80, while sensing of linker DNA by the Arp8 module ensures allosteric coupling of ATP hydrolysis to DNA translocation, which has been proposed to prevent back-slippage during DNA strain build up[43,47].

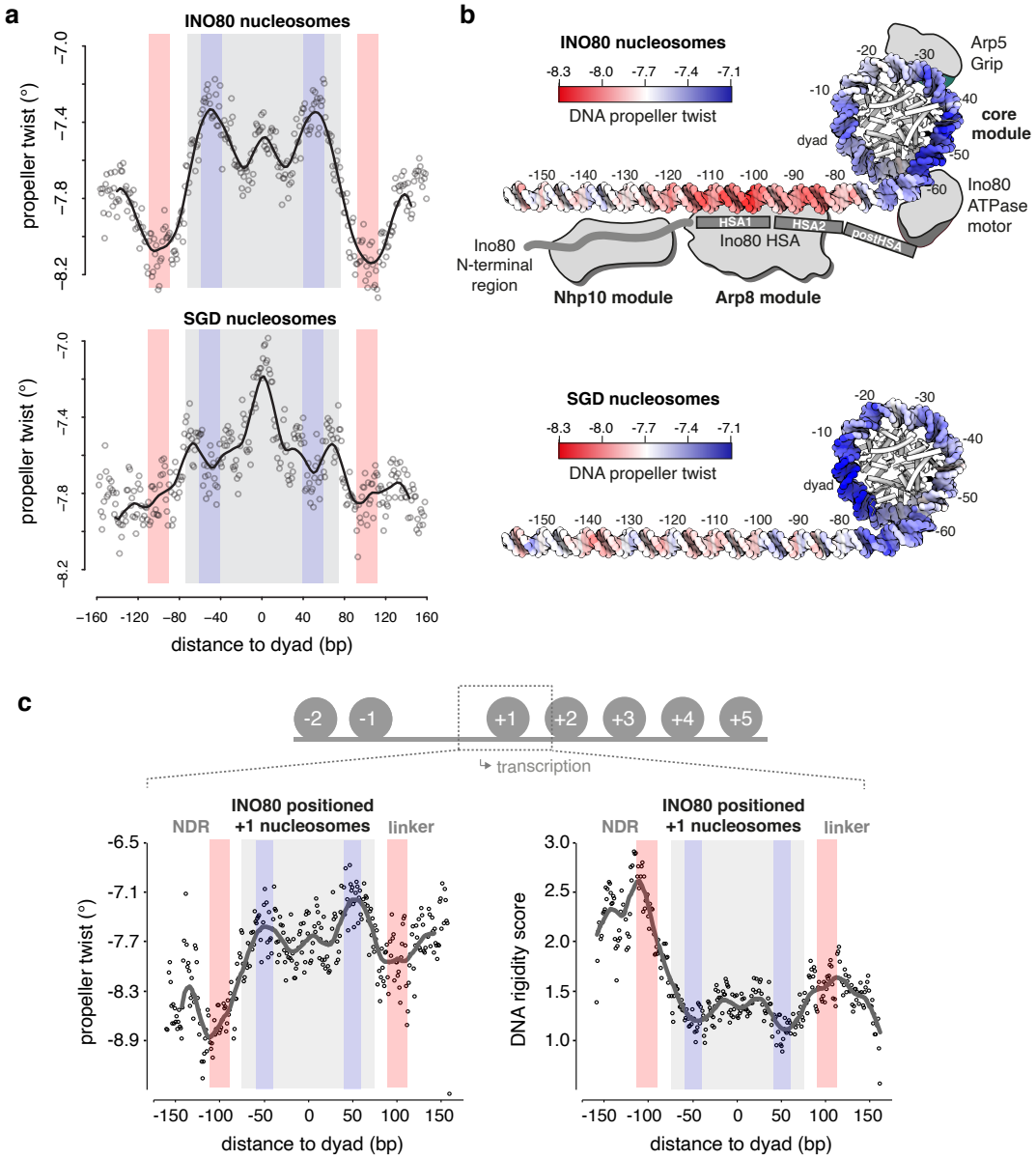


Figure 4. DNA shape readout underlies nucleosome positioning by INO80 and SGD. a Propeller twist DNA shape profiles for nucleosomal sequences of SGD chromatin with (INO80 nucleosomes) or without (SGD nucleosomes) remodeling by recombinant *S. cerevisiae* WT INO80 complex. Light red and light blue background indicate regions of major differences between SGD and INO80 profiles. Light grey background marks the location of the nucleosome core particle. **b** Red-white-blue color gradient mapping of propeller twist DNA shape profile from panel (a) on model of linker and nucleosomal core DNA. Binding architecture of INO80 is shown schematically and based on structural data[43,46] and biochemical mapping[47]. **c** Propeller twist DNA shape and DNA rigidity profiles for INO80 positioned +1 nucleosomes, all with the same orientation relative to the direction of transcription. See main text and Materials and Methods for a description of the DNA rigidity score. Note that the promoter NDR around -100 bp corresponds to a rigid DNA motif, while the score indicates an increased flexibility around -55 bp between the ATPase-motor and the Arp5-grip of INO80 (see panel b).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

This congruency immediately suggests a molecular mechanism by which an active readout not only through recognition of ground-state average DNA shape features, but also via ATP-hydrolysis driven perturbation of mechanical properties of DNA leads to the positioning of nucleosomes. The most immediate mechanical property of the double-helix is conformational flexibility. To assess this property on a genomic scale, we introduced a rigidity score that characterizes how rigid/flexible DNA is within a local region at bp resolution[33]. We considered A-tracts of consecutive ApA (TpT) or ApT bp steps as dominant factor in increasing rigidity due to strong stacking interactions combined with inter-bp hydrogen bonds in the major groove[32,53]. The rigidity score accounts for the length of A-tracts as longer runs of ApA (TpT) and ApT steps without TpA steps or G/C bp increase rigidity of a DNA fragment. We observed that DNA rigidity is correlated with DNA shape features, and the correlation remains at a consistent level across INO80 positioned nucleosomes (Supplementary Fig. 2a,b,c). This analysis reveals that +1 nucleosome positioning by INO80 involves placement of nucleosomes where DNA flexibility is increased at the -55 bp region between the ATPase motor and the Arp5 grip, while the promoter NDR region harbors a rigid DNA element where the Arp8-module is located (Fig. 4c). Intriguingly, a similarly rigid promoter DNA motif at the same distance in respect to the +1 nucleosome was also identified in a parallel study, where DNA mechanics were measured experimentally on a genomic scale via library-based DNA circularization assays[54].

Altered Ino80-HSA-helix-DNA contacts affect DNA shape/mechanics readout by INO80. To establish causality, we probed whether the INO80-DNA contacts and different histones would affect the readout of DNA shape/mechanics. Nucleosomes positioned by WT INO80 clustered together with those positioned by mutant complexes where mutations affected the Nhp10 module, i.e., the Ino80 N-terminus or Nhp10 module subunits including the Nhp10 HMG Box (Fig. 5a). This corroborated our results regarding nucleosome positioning in promoter regions (Fig. 2d-f) and ruled out a major role for the Nhp10 HMG box in DNA shape/mechanics readout by INO80. In contrast, all mutant complexes impaired in HSA helix-DNA contacts, either the HQ1 or HQ2 mutation and each also in combination with the HMGII mutations, generated distinct clusters of nucleosome positions (Fig. 5a). Overall shape/mechanics preferences were not much affected if endogenous fly versus recombinant human histones were used (Fig. 5b). This validated our use of fly histones for the comparisons among WT and mutant INO80 complexes in this approach.

In total, there were three major classes of nucleosome positions, those generated by i) SGD, ii) WT INO80/Nhp10 module mutant complexes or iii) HSA helix mutant complexes (Fig. 5a). To investigate the differences in DNA sequence preferences only between the INO80 complexes and at minimal contribution of neighboring nucleosomes, we clustered only the respective samples with low assembly degree SGD chromatin (Fig. 5c) and compared the resulting DNA shape/mechanics profiles of clusters with clearly different occupancies among the INO80 complexes, e.g., cluster 1 versus 3 (Fig. 5c, Supplementary Fig. 3). Propeller twist signal profiles clearly differed between clusters that contained nucleosome positions preferentially generated by the HSA helix-mutated INO80 versus WT or Nhp10 module mutated complexes. In particular, the $\pm$ 100 bp region of the linker DNA showed a distinct shift of the propeller twist signal by more than 20 bp between cluster 1 and 3 (Fig. 5d). As this is the region where the Ino80 HSA domain contacts DNA (Fig. 4b), these data directly showed that these HSA helix-DNA contacts contributed to the DNA shape/mechanics readout during nucleosome positioning. Moreover, additional changes of propeller twist signals within the nucleosomal DNA region provided, in context of Ino80 HSA mutations, evidence for the allosteric interplay between the Arp8- and the core module of INO80[46,47]. We conclude that INO80 positions nucleosomes via a readout of DNA shape/mechanics profiles. This information and its readout are distinct from known DNA sequence preferences of histones suggesting that remodelers play an active role in translating genomic information into nucleosome positions, i.e., determine nucleosome positions through their specific molecular mechanism of remodeling.

The DNA sequence-specific barrier Reb1 regulates nucleosome positioning by INO80. Having established that INO80 reads DNA shape/mechanics features and translates this information via specific modules into nucleosome positions, we asked next whether INO80 is also capable of processing nucleosome positioning information from DNA sequence-specific barrier factors (Fig. 1b). Reb1 is a GRF important for promoter nucleosome organization *in vivo*[26]. Sequence-specific GRFs serve, via an unknown mechanism, as nucleosome positioning alignment point for remodelers like ISW1a or ISW2[29].

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

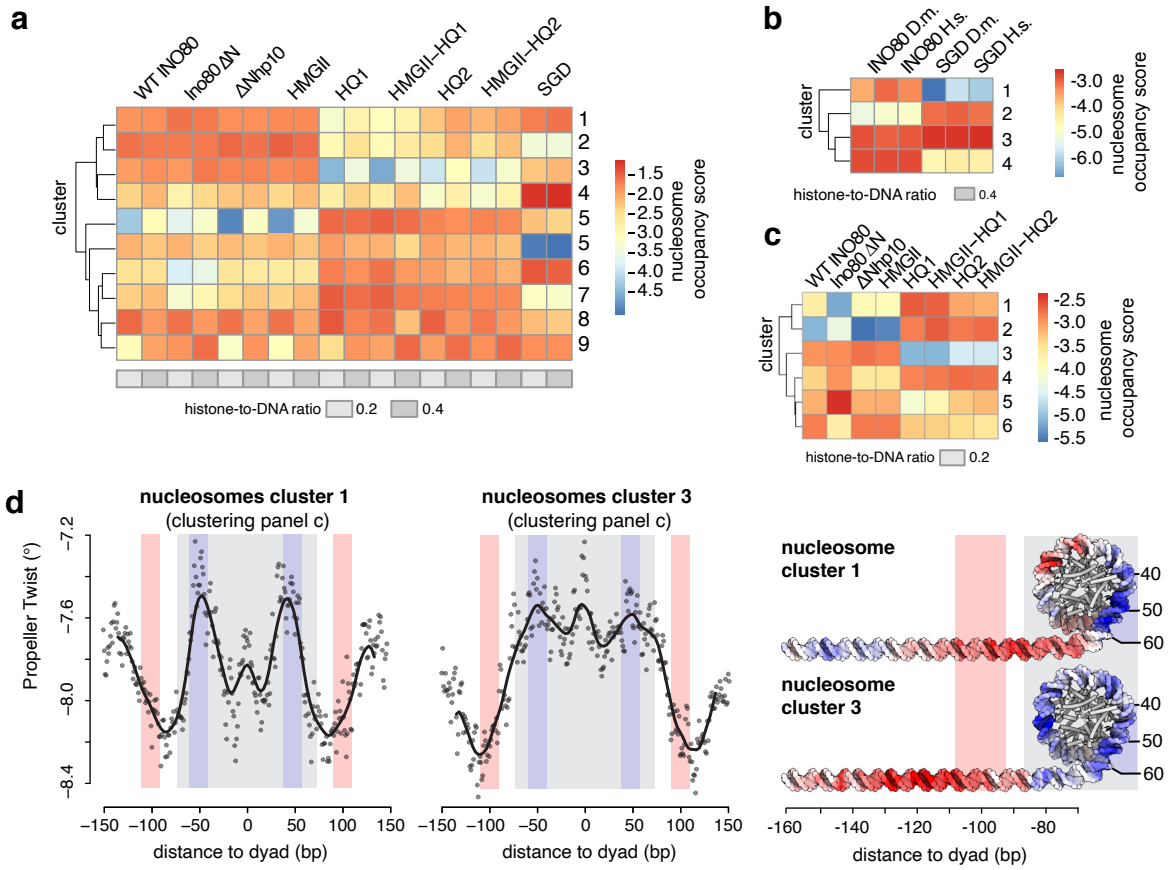


Figure 5. Structure-based mutations probe the DNA shape/mechanics readout by INO80. a Nucleosome position clusters derived from principal component analysis (PCA) of nucleosome positions of SGD chromatin with embryonic *D. melanogaster* histones at the indicated histone-to-DNA mass ratio without (SGD) or after remodeling by the indicated recombinant *S. cerevisiae* WT and mutant INO80 complexes (as in Figure 3d,e) **b** As panel a but for SGD chromatin with embryonic *D. melanogaster* (*D. m.*) vs. recombinant *H. sapiens* (*H. s.*) histones at the indicated histone-to-DNA mass ratio without (SGD) or with remodeling by recombinant *S. cerevisiae* WT INO80 complex (INO80). **c** As panel b but only with the indicated subset of samples. **d** As panel a but only for nucleosomes from the indicated clusters of panel c. Propeller twist DNA shape data mapped onto model of linker and nucleosomal DNA by using red-white-blue color gradient. See Supplementary Figure 3 for other clusters.

To directly test whether Reb1 binding at cognate promoter sites controls +1 nucleosome positioning by INO80, we turned again to whole-genome reconstitutions. Increasing Reb1 concentrations clearly improved nucleosome positioning by INO80 at promoters with Reb1 sites in terms of +1 nucleosome occupancy (peak height), but also in array extent and NDR depth (Fig. 6a,b, Supplementary Fig. 4a). This Reb1 effect was again independent of the histone octamer species-origin (Supplementary Fig. 4b). Detailed quantification of nucleosome spacing and array phasing at Reb1 sites and at different nucleosome densities was studied in the accompanying paper (Oberbeckmann & Niebauer et al.). *In vivo* mapping of INO80 subunits by ChIPexo[40] indicated that INO80 adopts an extended conformation, which might bridge Reb1 binding sites and +1 nucleosomes.

To directly address whether INO80 relays positioning information from Reb1 to +1 nucleosomes, we turned to classical mononucleosome assays. We generated mononucleosomes with a long linker DNA on one side from a promoter (of gene yGL167c) that was selected based on INO80 and Reb1 occupancy measured by ChIPexo *in vivo*[40] and clearly improved nucleosome positioning in whole-genome reconstitutions[29]. *In vivo*, the Reb1 site of the yGL167c promoter is 145 bp upstream of the +1 nucleosome dyad (about 72 bp to the 5' flank of the nucleosome core particle as the distance of this flank to the dyad is about 73 bp) which matches closely the median distance 149 ± 33 bp measured for all Reb1 sites at (median distance to the 5' flank of 76 ± 33 bp, Fig. 6f). We replaced the +1 nucleosome sequence by a Widom-601 nucleosome positioning sequence[55] and reconstituted with this construct (Fig. 6c, left) via SGD the *in vivo* promoter nucleosome architecture. Reb1 was added substoichiometrically to reconstituted yGL167c-NCP601 mononucleosomes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

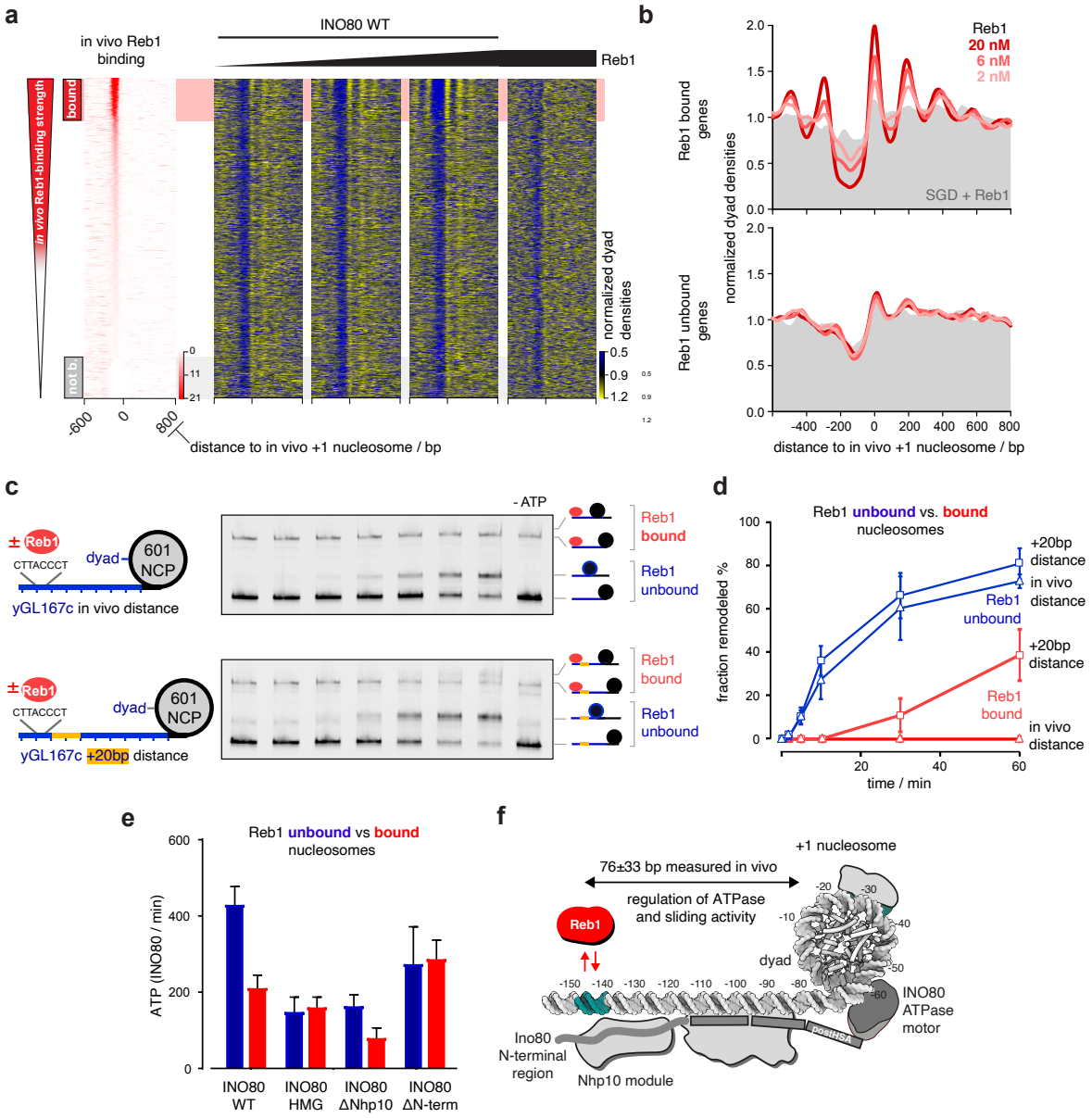


Figure 6. Reb1 regulates nucleosome positioning by INO80 and INO80's ATPase and sliding activity. a Heat maps of MNase-seq data for SGD chromatin assembled with recombinant *H. sapiens* histones at histone-to-DNA mass ratio 0.4, incubated with recombinant *S. cerevisiae* WT INO80 and increasing concentrations of recombinant Reb1 (ramp denotes 2, 6 and 20 nM Reb1). Right most panel shows sample prepared with embryonic *D. melanogaster* histones. Heat maps are aligned at *in vivo* +1 nucleosome positions and sorted according to decreasing (top to bottom) anti-Reb1 SLIM-ChIP score (*in vivo* Reb1 binding[68]) shown in leftmost heat map. Horizontal red or grey shading highlights genes with strong or weak *in vivo* Reb1 promotor-binding, respectively. Single replicates were plotted, see Supplementary Figure 3a for all replicates. **b** Composite plots of MNase-seq data as in panel A averaged over genes highlighted in red (top) or grey (bottom) in panel (a). **c** Left: mononucleosome substrate design with 80 bp (top) or 100 bp DNA overhang (bottom) taken from a promoter (yGL167c) with clear +1 nucleosome positioning by just INO80 *in vitro* and INO80 bound *in vivo* [40]. Guided by its dyad positions, we replaced the genomic +1 nucleosome sequence of yGL167c with a 601-nucleosome positioning sequence. Right: Native PAGE nucleosome sliding assay for indicated mononucleosome and Reb1 concentrations, and 10 nM recombinant *S. cerevisiae* WT INO80 for yGL167c-NCP601 (top) or yGL167c-20-NCP601 (bottom). "-ATP" denotes 60 min time point without ATP. **d** Quantification of sliding assays from the middle panel and two other replicates. Traces in red show data in the presence of Reb1. Error bar shows SD between replicates. **e** NADH-based ATPase assay for the 25 nM mononucleosomes and 10 nM recombinant *S. cerevisiae* WT and mutant INO80 complexes alone or with Reb1 at equimolar ratio to mononucleosome respectively. **f** Structural data[43,46] and biochemical mapping[47] suggest a putative binding architecture of INO80 which might bridge Reb1 and +1 nucleosomes. Allosteric communication occurs across a distance of more than 70 bp (median of 76±33bp *in vivo* measured by SLIM-ChIP[69] and MNase-seq).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

As separation in native polyacrylamide gel electrophoresis could distinguish mono-nucleosomes with and without bound Reb1, we could compare remodeling kinetics with and without Reb1 in the same reaction (Fig. 6c). Kinetics of sliding the initially end-positioned nucleosome to the center were much slower, if at all detectable, in the presence of Reb1 (Fig. 6c,d). As the distance between bound Reb1 and the 601-nucleosome was as *in vivo* and therefore, probably corresponded to the steady state distance set by INO80, we prepared and assayed in the same way a second construct (yGL167c-20-NCP601, Fig. 6c) with additional 20 bp of DNA inserted in the yGL167c promoter. This end-positioned 601-nucleosome was clearly moved towards the Reb1 barrier by INO80 (Fig. 6c), but again at a slower rate compared to sliding this nucleosome to the center in the absence of Reb1 (Fig. 6d).

We asked next, whether decreased sliding kinetics were caused by inhibition or by decoupling of ATPase activity. Notably, most INO80 mutations that abrogate nucleosome sliding, such as the HQ1/2 or Arp5 mutations, still showed robust ATPase activity[43,46]. In contrast, INO80 ATPase assays in the presence of yGL167c-NCP601 mononucleosomes showed about twofold decreased ATPase activity upon addition of Reb1 compared to reactions without Reb1 (Fig. 6e). This was not a general effect of Reb1 in this assay as the HMGII complex as well as the Ino80$^{\Delta N}$ INO80 mutant complexes with point mutations in the HMG box of Nhp10 or lacking the N-terminal region of Ino80 and the Nhp10-module, respectively (Fig. 2c, Supplementary Fig. 1h), did not show a reduction of ATPase activity upon Reb1 addition (Fig. 6e), while the ATPase activity of the ΔNhp10 INO80 mutant complex was still regulated by Reb1. The detailed mechanism of this intriguing allosteric communication across a distance of more than 70 bp linker DNA awaits further structural studies. However, based on the regulatory role of the N-terminal region of Ino80 even in the absence of the Nhp10 module, we cautiously speculate that it might serve not only as a binding platform for Nhp10, but that it stimulates the activity of INO80 in absence of Reb1 possibly via restricting the dynamics of the Arp8 module.

Taken together, we concluded that Reb1 binding to its cognate promoter sites regulates INO80 activity allosterically by inhibition through interaction via the N-terminal region of Ino80 that is modulated by the Nhp10 module subunits. The multi-subunit architecture of INO80 relays thereby positioning information between Reb1 and +1 nucleosomes, adjusts the +1 nucleosome to its *in vivo*-like position and programs thereby genic regions for formation of nucleosome arrays (Fig. 6f).

INO80 integrates information from DNA shape/mechanics and Reb1 at promoters. A Reb1 site at a distance to a nucleosome position corresponds to an input of DNA sequence information, mediated by its bound cognate factor Reb1, compared to the input of DNA shape/mechanics features. Therefore, we asked if and how INO80 serves as an information processing hub and integrates such different information input into resulting nucleosome positions.

First, we asked if promotors with Reb1 sites at all contained DNA shape information leading to +1 nucleosome positioning by INO80 on its own in the absence of Reb1. Maybe promoter regions had evolved such that +1 nucleosome positions were either directly encoded via DNA shape/mechanics or indirectly via GRF sites. We compared nucleosome positioning by INO80 in the absence of Reb1 at Reb1 site-containing promoters with positioning at an equal number of promoters lacking any GRF sites. As INO80 was able to position *in vivo*-like +1 nucleosomes on its own at both types of promoter regions (Fig. 7a), we concluded that both types contained +1 nucleosome position DNA shape/mechanics information in their genomic sequence.

Second, we asked if the additional information of bound Reb1 at the promoters with Reb1 site was synergistic, antagonistic or neutral to the DNA shape/mechanics-guided positioning by INO80. Comparing nucleosome positioning by INO80 at Reb1 site-containing promoters with versus without Reb1 showed that the Reb1 information mainly synergized with the DNA shape/mechanics information and led to very similar positions but, in keeping with the outcome of the Reb1 titration (Fig. 6a,b), to higher +1 nucleosome peaks and more pronounced NDRs (Fig. 7b). Quantification of the differences in resulting peak positions with vs. without Reb1 showed that +1 nucleosome peaks differed on average only by 6 ± 3 bp for SGD chromatin with histone-to-DNA mass ratios of 0.2 or 0.4, which was within the experimental error of our reconstitutions (Fig. 7c). For higher assembly degrees with a histone-to-DNA mass ratio of 0.8, the difference was 15 ± 5 bp, which was due to nucleosome positioning closer to Reb1 with increasing histone density, while the +1-nucleosome position as determined by INO80 on its own via DNA shape/mechanics was hardly affected by variations in nucleosome density. Nonetheless, high density affected peak heights, which is discussed, together with the effects of density on nucleosome distance

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

to barrier, in the accompanying paper (Oberbeckmann & Niebauer et al.) in the context of our remodeler ruler concept. Here, we concluded that genome sequence evolved a DNA shape/mechanics signal downstream of a Reb1 site in direction of transcription so that nucleosome positioning by INO80 either guided by DNA shape/mechanics or by Reb1 leads to very similar +1 nucleosome positions at low or medium nucleosome density. Note that promoter Reb1 sites are situated *in vivo* within NDRs[56], which, by definition, represent regions of locally low nucleosome density.

Third, we noted that the synergism between DNA shape/mechanics- and Reb1-guided nucleosome positioning by INO80 only applied to the +1 nucleosome in direction of transcription, but not to the -1 nucleosome, as we observed in our reconstitution experiments in in vivo-like differences between the respective MNase-seq peak heights (Fig. 7b).

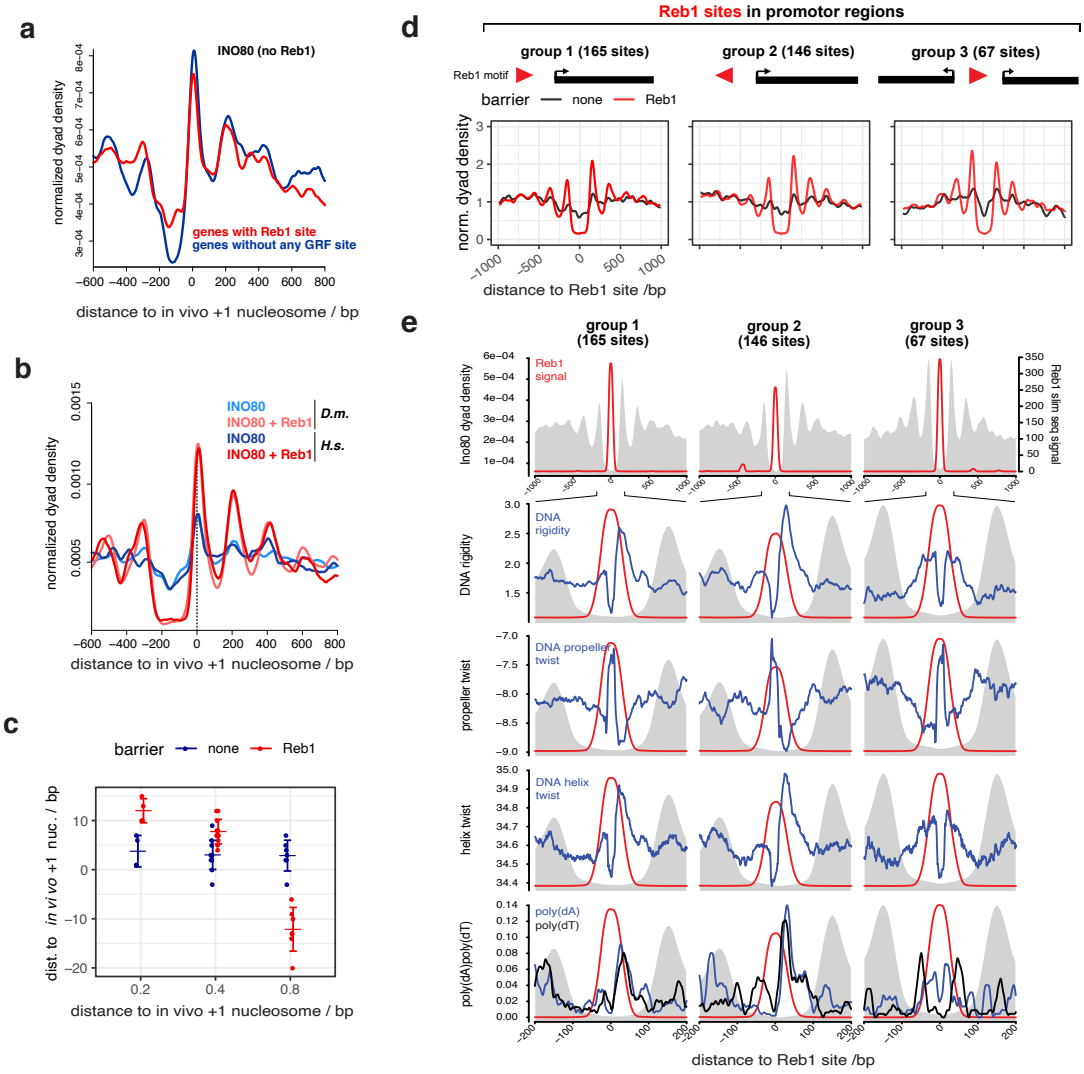


Figure 7. INO80 synergistically integrates nucleosome positioning information from DNA shape and Reb1 barriers. a Composite plots as in Figure 6b but for SGD chromatin with recombinant human histones at 0.4 histone-to-DNA mass ratio incubated with recombinant *S. cerevisiae* WT INO80 plotted for either genes with promoter Reb1 sites only (as red shading in Figure 6a) or for a randomly selected but similar number of genes with no GRF sites (Abf1, Rap1, Mcm1, Cbf1 [70]) in their promoters. **b** As panel a but for merged replicates comparing SGD chromatin with embryonic fly (*D. m.*) or recombinant human (*H. s.*) histones, ± 20 nM Reb1 and only for genes with promoter Reb1 sites. **c** Distributions of distances between +1 nucleosome positions at Reb1-site containing promoters reconstituted by incubation of SGD chromatin with the indicated histone-to-DNA mass ratio with recombinant *S. cerevisiae* WT INO80 in the presence (Reb1) or absence (none) of 20 nM Reb1. **d** As Figure 6b, but aligned at Reb1 sites of the indicated groups and with recombinant *S. cerevisiae* WT INO80 ± 20 nM Reb1. **e** Reb1 site-aligned composite plots for genes groups as in panel d, from top to bottom: positions of Reb1 site PWM motifs, Reb1 site motifs and DNA rigidity, Reb1 sites and propeller twist DNA shape features and Reb1 motifs and positions of poly(dA) or poly(dT) elements (> 6 homopolymeric stretches). Grey background in all panels shows composite plot of MNase-seq data as in panel d.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To assess this point more clearly and to ask if orientation of the intrinsically asymmetric Reb1 site further affected nucleosome positioning, we grouped Reb1 site-containing promoters according to the Reb1 site orientation relative to neighboring genes (groups 1 to 3, Fig. 7d). Reb1 site-aligned MNase-seq data composite plots averaged over genes within these groups showed that peak heights and array generation were more pronounced in direction of transcription but independent of Reb1 site orientation. This further supported our conclusion that synergistic DNA shape/mechanics information evolved next to Reb1 sites only in places where a +1 nucleosome becomes positioned that plays the well-known role in regulation of transcription initiation[4,28]. Accordingly, promoters in groups 1 to 3 showed distinct asymmetrical DNA shape/mechanics features and strand-specific poly(dA:dT) prevalence in the direction of transcription (Fig. 7e). Thus, these data suggest that INO80-mediated +1 nucleosome positioning is symmetrically guided by Reb1 as orientation of the Reb1 site did not matter (group 1 vs. 2, Fig. 7d). Importantly, however, our analysis revealed that Reb1 sites at promoters evolved synergistically with DNA shape/mechanics features, which explains the observed peak height asymmetry (groups 1 and 2) or symmetry (group 3) of nucleosome patterns depending on the DNA shape/mechanics feature distribution in the genome (Fig. 7e). The deviations in +1 nucleosome positions between DNA shape/mechanics- versus Reb1-guided positioning (Fig. 7c) in response to nucleosome density suggest that Reb1-guided positioning is either dominant or that Reb1-guided positioning is still equally effective at high density while DNA shape/mechanics-guided positioning is impaired. In the accompanying paper (Oberbeckmann & Niebauer et al.) we show that the latter is the case.

DNA ends are potent barriers for INO80 nucleosome positioning. Having established a synergy between DNA shape/mechanics and Reb1 sites at gene promoter regions, we asked whether we can uncouple barrier-mediated positioning from a promoter sequence context. To test this idea, we analyzed nucleosome positioning at all *in vivo* mapped genomic Reb1 sites (Fig. 8a,b). Consistent with our findings above, we observed symmetrical nucleosome arrays around all Reb1 sites (Fig. 8b, top right) suggesting that barrier-mediated positioning can occur independently of other DNA sequence features. In light of this, we considered that INO80 may align nucleosomes also to different barrier types as long as they represented a clear alignment point. In our search of the minimalistic system that provides nucleosome positioning information, we wondered if simply a DNA end could constitute a barrier. Notably, INO80 has been involved in DNA damage response signaling upon DNA double strand breaks (DSBs) *in vivo*[57]. In principle, such as scenario was already tested in classical mononucleosome sliding assays as these automatically involve two DNA ends.

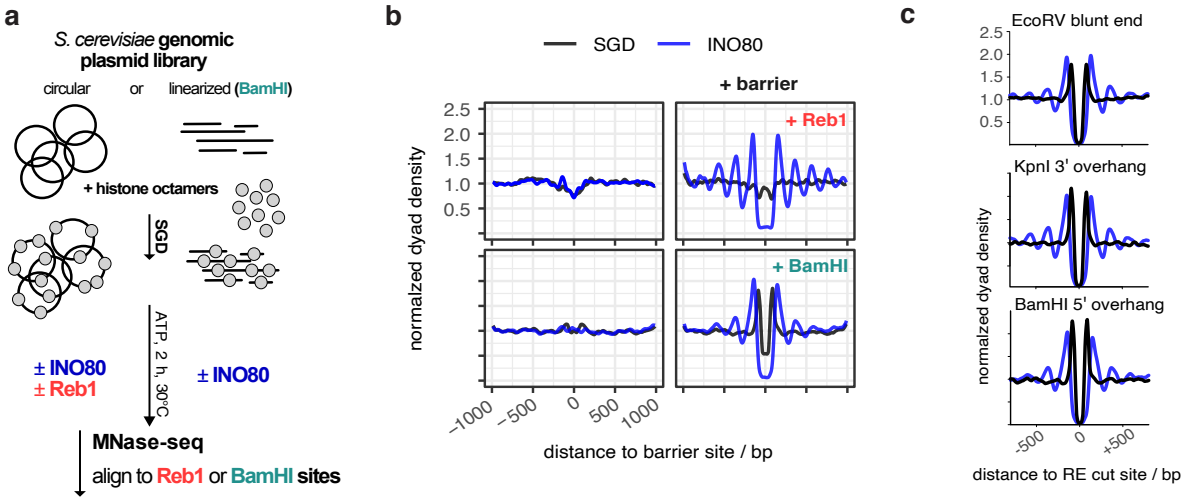


Figure 8. DNA ends are potent barriers for nucleosome positioning by INO80. a Overview (analogous to Figure 2a) of reconstitution with circular versus RE-precleaved plasmid libraries. **b** Composite plots of BamHI-site aligned versus anti-Reb1 SLIM-ChIP-defined Reb1 sites aligned MNase-seq data for: top, SGD prepared with circular plasmid library and incubated without (SGD) or with recombinant *S. cerevisiae* WT INO80 (INO80), and bottom: as top but with BamHI-precleaved library if indicated (+ BamHI). **c** As panel b, but for SGD chromatin with plasmid libraries pre-cleaved with the indicated RE and data aligned at the indicated RE cut sites. Strong peaks flanking cut RE sites in SGD chromatin without INO80 remodeling reflected an MNase-seq bias. Due to the pre-cleavage, the probability that a mono-nucleosomal fragment flanking the cut site is released by MNase is increased.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

However, effects there may have been due to the comparatively short length of template DNA and to the presence of two DNA ends at the same time. Our genome-wide system allowed us to test the effect of one-sided DNA ends in the context of very long DNA. We introduced double stranded DNA ends at fortuitous locations, i.e., without likely evolutionarily shaped context, throughout the *S. cerevisiae* genome via restriction enzyme (RE) digest of the plasmid library prior to SGD reconstitution (Fig. 8a). As expected, SGD chromatin neither with nor without remodeling by INO80 showed distinct nucleosome patterns at uncleaved BamHI sites (Fig. 8b, bottom left). However, strong and symmetrical arrays were aligned at cut sites by INO80 (Fig. 8b, bottom right). The same was true for other REs that generated different kinds of DNA ends (Fig. 8c). We concluded that all three kinds of DNA ends (blunt, 3' or 5' overhang) were strong nucleosome positioning barriers for INO80.

Discussion

In this study, we identified and probed the fundamental molecular determinants by which ATP-dependent chromatin remodelers position nucleosomes across the genome. An integrated approach combining fully-recombinant, *de novo* whole-genome reconstitutions, high-resolution structural information, and PCA/clustering analysis revealed that the INO80 complex processes DNA sequence information, both via readout of a distinct DNA shape/mechanics signature motif, as well as, via alignment against a DNA sequence specific barrier factor like Reb1 or at DSBs. INO80's multi-subunit architecture integrates the readout of different positioning information, contributes through its mechanism its own information and determines thereby how this is translated into positions of +1 and other nucleosomes (Fig. 9).

Although the pivotal role of remodelers in chromatin organization and their dependency on DNA sequences has been recognized[29,31,58], nucleosome positioning sequences (NPSs) were usually defined as sequences of "intrinsic" positioning by SGD driven solely by histone octamer-DNA interactions, as illustrated by the Widom-601 NPS[55]. PCA/clustering analysis enabled us now to reassess these classical SGD-NPSs and to identify a new kind of NPS. We find that SGD-NPSs correspond to distinct DNA-sequence dependent shape/mechanics profiles, while nucleosome positioning by a remodeler like INO80 corresponds to a different shape/mechanics profile. Therefore, we identified the latter as INO80-NPSs.

Respective remodeler-NPSs are likely to exist for other remodelers and it will be interesting where they evolved in genomes. The mere observation that INO80 and RSC remodelers generate different nucleosome positions, despite working on the same histone octamers and DNA sequences, suggested previously[29, 59] that remodelers do not just allow histone octamers to occupy their thermodynamically preferred positions (otherwise different remodelers would generate the same positions), but that remodelers, as demonstrated in this study, read genomic information, actively override octamer preferences and shape the positioning landscape in a remodeler-specific way. In analogy to the "genomic code for nucleosome positioning", i.e. the proposed evolution of SGD-NPSs, evolved remodeler-NPSs would implement a "remodeler code for nucleosome positioning" as proposed earlier[59]. We abstain from adding another "code" to the troubled epigenetics discussions but point out the conceptual analogy.

Importantly, we go here beyond a mere correlation between INO80-NPSs and DNA shape/mechanics profiles. The causal mechanistic link was directly established by tuning the INO80 DNA shape/mechanics readout via targeted INO80 mutations. Informed by high-resolution structures, we found independently that on the one hand mutation of Ino80-HSA-DNA contacts more than ~100 bp away from the nucleosome dyad caused altered nucleosome positioning patterns, while on the other hand unbiased PCA/clustering analysis revealed also altered DNA shape/mechanics features right in the same region. Together, our results provide strong evidence for a readout of these DNA shape/mechanics features. Moreover, we observed altered processing of DNA shape/mechanics features at the -55 bp region between the Ino80 core ATPase motor and the Arp5 grip, suggesting a critical role of DNA shape/mechanics in regulating the build-up of DNA strain during the core mechanism of nucleosome translocation[43,47,48]. Intriguingly, the effects at both regions are coupled via two allosteric communication pathways of possibly equal importance: on the protein side, linker DNA recognition by the Arp8 module is coupled to the activity of the Ino80 ATPase motor of the core module via the extended helical configuration of the HSA and postHSA domains[46]. On the DNA side, DNA shape/mechanics features at the histone-bound -55 bp region are most likely coupled to DNA shape/mechanics features at the DNA linker -100 bp region in the context of over- and underwinding of DNA in front and behind the Ino80 ATPase motor

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[38,43]. More generally, our data illustrates a regulatory circuitry comprising a two-way relationship between a protein factor working on DNA and DNA properties feeding back to the protein factor. Overall, INO80-NPSs represent the nucleosome positioning information that emerges from the combination of DNA, histones, and the active interpretation via the allosteric communication within the remodeler.

For these reasons, the DNA shape/mechanics readout by INO80 importantly expands the scope of recently discussed DNA shape contributions. DNA shape was mostly studied in the context of "static" DNA binding, e.g., by transription factors and GRFs[60-62]. In contrast, INO80 dynamically reads and interprets DNA shape/mechanics while tracking along DNA in an ATP-dependent manner. Thereby, INO80 actively probes the mechanical properties of DNA. Thus, this read out of genome information is expected to serve as a role model for other factors that translocate along DNA or also RNA, like other remodelers, helicases, cohesins or polymerases. For example, RNA polymerase I was suggested to read the DNA bend at its promoters[63] and RNA polymerase II may recognize its promoters via structural DNA features (bending, meltability, flexibility) rather than via classical consensus sequences[64]. As these structural properties are redundantly linked to DNA sequence, we propose that readout of such DNA structural properties may be common if factors deal with a wide range of genomic regions.

As alternative DNA sequence signals, there is DNA sequence information of classical consensus motifs for specific binding by cognate factors. GRFs are well-known to program +1 nucleosome positioning and formation of genic nucleosome arrays *in vivo*[26,34,65]. In light of our finding that DNA ends are also potent nucleosome positioning barriers, it is tempting to speculate that remodelers involved in DNA damage response, such as INO80[57], may generate regular nucleosome arrays as a licensing platform at DSBs *in vivo*.

The mechanism by which remodelers generate arrays at barriers, i.e., read positioning information via an alignment mechanism, remained largely unknown. This study reveals that nucleosome positioning by INO80 is actively regulated by Reb1 at promoter sites through an interaction with the N-terminal region of Ino80 (Fig. 9a). Intriguingly, Reb1 decreased not only nucleosome sliding, but also inhibited ATPase activity of INO80, even at a distance of -145 bp between the cognate Reb1 site and the dyad of the +1 nucleosome. In contrast, DNA linker length sensing by INO80 at DNA ends uncouples a decrease in mononucleosome sliding from its robust stimulation of ATPase activity[46,48]. Consequently, GRFs might represent a different kind of regulatory barrier compared to DSBs, at least in the absence of the DNA repair machinery. In the accompanying study (Oberbeckmann & Krietenstein et al.), we identify the Arp8-module and the Nhp10 module as a multi-layered ruler element which measures and sets nucleosome arrays differently in respect to Reb1 sites, DNA ends and neighboring nucleosome. Taken together, our findings lead to a model how regulation of nucleosome sliding direction bias upon interaction with a barrier can lead to stable nucleosome positioning and array formation. The multi-subunit architecture of INO80 functions similarly to a relay: INO80 receives input via its Arp8 and Nhp10 modules and communicates this information allosterically towards the ATPase of the INO80 core, where it is translated into a nucleosome position (Fig. 9b).

The exact +1 nucleosome position impacts transcription regulation, e.g., it differs between repressed and activated promoters and influences TSS selection[4,11,28,66]. In this study, we show that these positions are robustly encoded in the genome in two ways, i.e., both by DNA shape/mechanics features and corresponding distances to the Reb1 site. Nucleosome positioning next to Reb1 did not require DNA shape/mechanics features as it also worked symmetrically on the other side even if there was no evolved promoter. Importantly, however, in context of promoter regions, we identify a co-evolved synergy between DNA shape/mechanics signatures and Reb1 binding sites, leading to asymmetric +1 nucleosome positioning, as measured by MNase-seq peak heights. This synergy provides not only robustness, but also an inroad to regulation. For example, we show that Reb1-mediated positioning is altered in response to nucleosome densities. Thus, we propose that regulation of nucleosome density at promoters, e.g., via the local activity of RSC, the major nucleosome-evicting remodeler in yeast[23], may result in regulation of +1 nucleosome positions. With high RSC activity, local promoter nucleosome density is low and +1 nucleosome positioning by INO80 coincides for DNA shape/mechanics- and Reb1-information input. Upon low RSC activity, nucleosome density is high and INO80 disregards the shape/mechanics signal and places the +1 nucleosome closer to Reb1, which corresponds to the more upstream +1 nucleosome position implicated in repressed promoter states.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

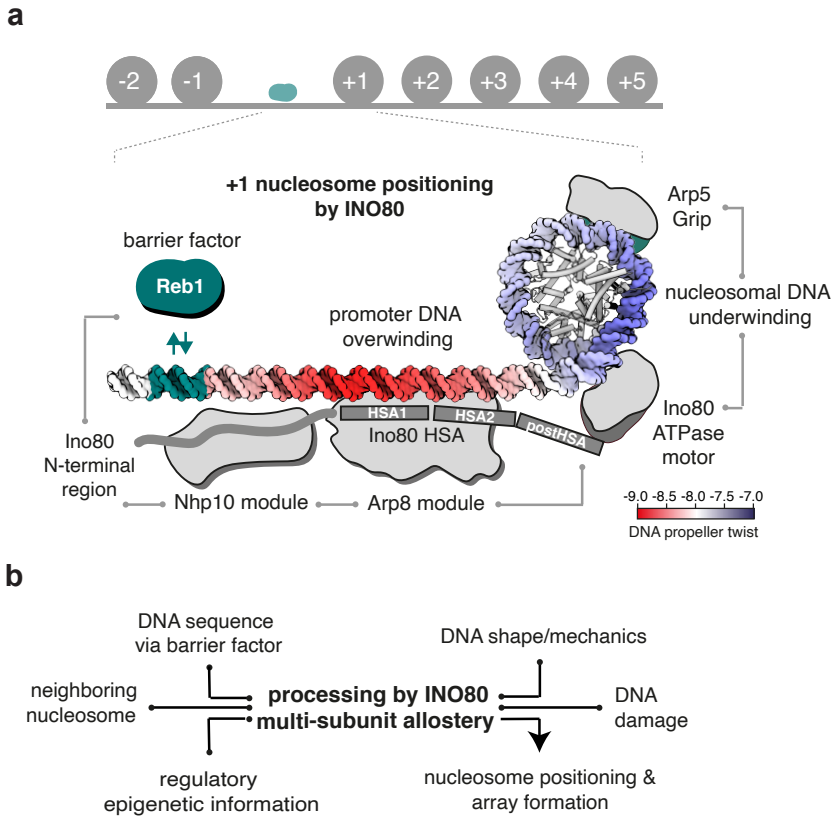


Figure 9. Model of +1 nucleosome positioning by INO80. a INO80 synergistically processes genomic information derived from DNA shape/mechanics as well as DNA sequence motifs bound by GRF Reb1 to position +1 nucleosomes. Structural data[43,46], biochemical[47] and ChIP-exo mapping[40] suggest a binding architecture of INO80 at +1 nucleosomes that is fully consistent with the identified positioning information and mechanism. Promoter DNA overwinding and nucleosomal DNA underwinding is derived from the direction of DNA translocation by the Snf2-type ATPase of INO80[43]. Allosteric communication is indicated by grey lines. **b** Signal integration and processing by multi-subunit allostery within INO80 leads to nucleosome positioning and array formation. Epigenetic information such as histone marks are expected to provide an additional layer of regulatory input, e.g., in response to the physiological state of the cell.

By genome wide biochemistry, this study reveals that a minimal set of information, comprising genomic DNA sequences, globular histones, and the molecular machinery of the remodeler, is sufficient to explain the placement and regulation of nucleosomes at their *in vivo* +1 positions for many promoters where appropriate DNA shape/mechanics signatures evolved. The identified mechanism of active information processing (Fig. 9b) provides allosteric control and versatile means for selective regulation, e.g., by epigenetic information such as histone modifications and variants as well as by the presence of sequence-specific factors such as transcription factors and pioneer factors. Signal integration of genome information from DNA shape/mechanics and sequence specified GRF binding by the multi-subunit architecture of INO80 exemplifies such principles. In the accompanying paper (Oberbeckmann & Niebauer et al.), we show how information from GRFs, DNA ends and positioned nucleosomes can be propagated into regular nucleosome arrays and how this process is regulated by remodeler rulers and nucleosome density. Collectively, this makes ATP dependent remodelers the fundamental information processing hub for nucleosome positioning and thereby the primary architects of the first level of chromatin organization.

References

1. Bell, O., Tiwari, V.K., Thoma, N.H. & Schubeler, D. Determinants and dynamics of genome accessibility. *Nat Rev Genet* **12**, 554-64 (2011).
2. Lai, W.K.M. & Pugh, B.F. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat Rev Mol Cell Biol* **18**, 548-562 (2017).
3. Bannister, A.J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381-95 (2011).
4. Jiang, C. & Pugh, B.F. Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* **10**, 161-72 (2009).
5. Lawrence, M., Daujat, S. & Schneider, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet* **32**, 42-56 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

6. Mavrich, T.N. et al. Nucleosome organization in the Drosophila genome. *Nature* **453**, 358-62 (2008).
7. Yuan, G.C. et al. Genome-scale identification of nucleosome positions in S. cerevisiae. *Science* **309**, 626-30 (2005).
8. Zink, L.M. & Hake, S.B. Histone variants: nuclear function and disease. *Curr Opin Genet Dev* **37**, 82-89 (2016).
9. Lieleg, C., Krietenstein, N., Walker, M. & Korber, P. Nucleosome positioning in yeasts: methods, maps, and mechanisms. *Chromosoma* **124**, 131-51 (2015).
10. Haberle, V. & Lenhard, B. Promoter architectures and developmental gene regulation. *Semin Cell Dev Biol* **57**, 11-23 (2016).
11. Shivaswamy, S. & Iyer, V.R. Stress-dependent dynamics of global chromatin remodeling in yeast: dual role for SWI/SNF in the heat shock stress response. *Mol Cell Biol* **28**, 2221-34 (2008).
12. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-60 (1997).
13. Widom, J. Role of DNA sequence in nucleosome stability and dynamics. *Q Rev Biophys* **34**, 269-324 (2001).
14. Satchwell, S.C., Drew, H.R. & Travers, A.A. Sequence periodicities in chicken nucleosome core DNA. *J Mol Biol* **191**, 659-75 (1986).
15. Kaplan, N. et al. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362-6 (2009).
16. Kornberg, R.D. & Stryer, L. Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism. *Nucleic Acids Res* **16**, 6677-90 (1988).
17. Zhang, Y. et al. Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat Struct Mol Biol* **16**, 847-52 (2009).
18. Zhang, Z. et al. A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* **332**, 977-80 (2011).
19. Gossett, A.J. & Lieb, J.D. In vivo effects of histone H3 depletion on nucleosome occupancy and position in Saccharomyces cerevisiae. *PLoS Genet* **8**, e1002771 (2012).
20. van Bakel, H. et al. A compendium of nucleosome and transcript profiles reveals determinants of chromatin architecture and transcription. *PLoS Genet* **9**, e1003479 (2013).
21. Lieleg, C. et al. Nucleosome spacing generated by ISWI and CHD1 remodelers is constant regardless of nucleosome density. *Mol Cell Biol* **35**, 1588-605 (2015).
22. Mobius, W. & Gerland, U. Quantitative test of the barrier nucleosome model for statistical positioning of nucleosomes up- and downstream of transcription start sites. *PLoS Comput Biol* **6**(2010).
23. 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).
24. Gkikopoulos, T. et al. A role for Snf2-related nucleosome-spacing enzymes in genome-wide nucleosome organization. *Science* **333**, 1758-60 (2011).
25. Cairns, B.R. et al. RSC, an essential, abundant chromatin-remodeling complex. *Cell* **87**, 1249-60 (1996).
26. Hartley, P.D. & Madhani, H.D. Mechanisms that specify promoter nucleosome location and identity. *Cell* **137**, 445-58 (2009).
27. Kubik, S. et al. Sequence-Directed Action of RSC Remodeler and General Regulatory Factors Modulates +1 Nucleosome Position to Facilitate Transcription. *Mol Cell* **71**, 89-102.e5 (2018).
28. Kubik, S. et al. Opposing chromatin remodelers control transcription initiation frequency and start site selection. *Nat Struct Mol Biol* **26**, 744-754 (2019).
29. Krietenstein, N. et al. Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell* **167**, 709-721.e12 (2016).
30. Krietenstein, N., Wippo, C.J., Lieleg, C. & Korber, P. Genome-wide in vitro reconstitution of yeast chromatin with in vivo-like nucleosome positioning. *Methods Enzymol* **513**, 205-32 (2012).
31. Lorch, Y., Maier-Davis, B. & Kornberg, R.D. Role of DNA sequence in chromatin remodeling and the formation of nucleosome-free regions. *Genes Dev* **28**, 2492-7 (2014).
32. Rohs, R., West, S.M., Liu, P. & Honig, B. Nuance in the double-helix and its role in protein-DNA recognition. *Current Opinion in Structural Biology* **19**, 171-177 (2009).
33. Slattery, M. et al. Absence of a simple code: how transcription factors read the genome. *Trends in Biochemical Sciences* **39**, 381-399 (2014).
34. Challal, D. et al. General Regulatory Factors Control the Fidelity of Transcription by Restricting Non-coding and Ectopic Initiation. *Mol Cell* **72**, 955-969.e7 (2018).
35. Fourel, G., Miyake, T., Defossez, P.A., Li, R. & Gilson, E. General regulatory factors (GRFs) as genome partitioners. *J Biol Chem* **277**, 41736-43 (2002).
36. Tsankov, A., Yanagisawa, Y., Rhind, N., Regev, A. & Rando, O.J. Evolutionary divergence of intrinsic and trans-regulated nucleosome positioning sequences reveals plastic rules for chromatin organization. *Genome Res* **21**, 1851-62 (2011).
37. Yamada, K. et al. Structure and mechanism of the chromatin remodelling factor ISW1a. *Nature* **472**, 448-53 (2011).
38. Jungblut, A., Hopfner, K.-P. & Eustermann, S. Megadalton chromatin remodelers: common principles for versatile functions. *Current Opinion in Structural Biology* **64**, 134-144 (2020).
39. Jones, G.M. et al. A systematic library for comprehensive overexpression screens in Saccharomyces cerevisiae. *Nature Methods* **5**, 239-241 (2008).
40. Yen, K., Vinayachandran, V. & Pugh, B.F. SWR-C and INO80 chromatin remodelers recognize nucleosome-free regions near +1 nucleosomes. *Cell* **154**, 1246-56 (2013).
41. Bonisch, C. & Hake, S.B. Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res* **40**, 10719-41 (2012).
42. White, C.L., Suto, R.K. & Luger, K. Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *Embo j* **20**, 5207-18 (2001).
43. Eustermann, S. et al. Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* **556**, 386-390 (2018).
44. Udugama, M., Sabri, A. & Bartholomew, B. The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. *Mol Cell Biol* **31**, 662-73 (2011).
45. Ayala, R. et al. Structure and regulation of the human INO80-nucleosome complex. *Nature* **556**, 391-395 (2018).
46. Knoll, K.R. et al. The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling. *Nat Struct Mol Biol* **25**, 823-832 (2018).
47. Brahma, S., Ngubo, M., Paul, S., Udugama, M. & Bartholomew, B. The Arp8 and Arp4 module acts as a DNA sensor controlling INO80 chromatin remodeling. *Nat Commun* **9**, 3309 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

48. Zhou, C.Y. et al. The Yeast INO80 Complex Operates as a Tunable DNA Length-Sensitive Switch to Regulate Nucleosome Sliding. *Mol Cell* **69**, 677-688.e9 (2018).
49. Zhou, C.Y. et al. Regulation of Rvb1/Rvb2 by a Domain within the INO80 Chromatin Remodeling Complex Implicates the Yeast Rvbs as Protein Assembly Chaperones. *Cell Rep* **19**, 2033-2044 (2017).
50. Klein-Brill, A., Joseph-Strauss, D., Appleboim, A. & Friedman, N. Dynamics of Chromatin and Transcription during Transient Depletion of the RSC Chromatin Remodeling Complex. *Cell Reports* **26**, 279-+ (2019).
51. Mendiratta, S., Bhatia, S., Jain, S., Kaur, T. & Brahmachari, V. Interaction of the Chromatin Remodeling Protein hINO80 with DNA. *PLoS One* **11**, e0159370 (2016).
52. Zhou, T. et al. DNAshape: a method for the high-throughput prediction of DNA structural features on a genomic scale. *Nucleic Acids Res* **41**, W56-62 (2013).
53. Nelson, H.C.M., Finch, J.T., Luisi, B.F. & Klug, A. The Structure of an Oligo(Da).Oligo(Dt) Tract and Its Biological Implications. *Nature* **330**, 221-226 (1987).
54. Basu, A. et al. Measuring DNA mechanics on the genome scale. 2020.08.17.255042 (2020).
55. 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).
56. Rhee, H.S. & Pugh, B.F. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* **147**, 1408-19 (2011).
57. van Attikum, H., Fritsch, O., Hohn, B. & Gasser, S.M. Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair. *Cell* **119**, 777-88 (2004).
58. Winger, J. & Bowman, G.D. The Sequence of Nucleosomal DNA Modulates Sliding by the Chd1 Chromatin Remodeler. *J Mol Biol* **429**, 808-822 (2017).
59. Rippe, K. et al. DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes. *Proc Natl Acad Sci U S A* **104**, 15635-40 (2007).
60. Levo, M. et al. Unraveling determinants of transcription factor binding outside the core binding site. *Genome Res* **25**, 1018-29 (2015).
61. Rossi, M.J., Lai, W.K.M. & Pugh, B.F. Genome-wide determinants of sequence-specific DNA binding of general regulatory factors. *Genome Res* **28**, 497-508 (2018).
62. Zentner, G.E., Kasinathan, S., Xin, B., Rohs, R. & Henikoff, S. ChEC-seq kinetics discriminates transcription factor binding sites by DNA sequence and shape in vivo. *Nat Commun* **6**, 8733 (2015).
63. Engel, C. et al. Structural Basis of RNA Polymerase I Transcription Initiation. *Cell* **169**, 120-131.e22 (2017).
64. Dienemann, C., Schwalb, B., Schilbach, S. & Cramer, P. Promoter Distortion and Opening in the RNA Polymerase II Cleft. *Mol Cell* **73**, 97-106.e4 (2019).
65. Ganapathi, M. et al. Extensive role of the general regulatory factors, Abf1 and Rap1, in determining genome-wide chromatin structure in budding yeast. *Nucleic Acids Res* **39**, 2032-44 (2011).
66. Boeger, H., Griesenbeck, J., Strattan, J.S. & Kornberg, R.D. Nucleosomes unfold completely at a transcriptionally active promoter. *Mol Cell* **11**, 1587-98 (2003).
67. Ocampo, J., Chereji, R.V., Eriksson, P.R. & Clark, D.J. The ISW1 and CHD1 ATP-dependent chromatin remodelers compete to set nucleosome spacing in vivo. *Nucleic Acids Res* **44**, 4625-35 (2016).
68. Gutin, J. et al. Fine-Resolution Mapping of TF Binding and Chromatin Interactions. *Cell Rep* **22**, 2797-2807 (2018).
69. Gutin, J. et al. Fine-Resolution Mapping of TF Binding and Chromatin Interactions. *Cell Reports* **22**, 2797-2807 (2018).
70. Yan, C., Chen, H. & Bai, L. Systematic Study of Nucleosome-Displacing Factors in Budding Yeast. *Mol Cell* **71**, 294-305.e4 (2018).

Acknowledgments

We thank Stefan Krebs and Helmut Blum at the Laboratory for Functional Genome Analysis (LAFUGA, Gene Center, LMU München) for high throughput sequencing, and Sigurd Braun for access to and help with the plating robot. We thank Marianne Schwarz and Jens Michaelis for helpful discussions and earlier experiments using recombinant histones and endogenously purified INO80 not reported in this paper. This study was funded by the German Research Foundation (SFB1064 to P.K. and K.P.H. and Gottfried Wilhelm Leibniz-Prize to K.P.H.), the European Research Council (ERC Advanced Grant "INO3D" to K.P.H.), and the NIH (grant R35GM130376 to R.R.) and HFSP (grant RGP0021/2018 to R.R.). N.K. is supported by HFSP grant LT000631/2017-L.

Author contributions

Conceptualization: EO, NK, TS, RR, KPH, PK, SE; Data curation: EO, NK, VN; Formal analysis: EO, NK, VN, YW, TS; Funding acquisition, Project administration, Supervision: KPH, PK, SE; Investigation: EO, NK, VN, KS, MM, YW, TS, SE; Methodology: EO, NK, VN, KS, RR, TS, PK, SE; Validation: EO, NK, VN, KS, YW, RR, TS, PK, SE; Visualization: EO, NK, VN, YW, TS, PK, SE; Writing original draft: EO, NK, PK, SE; Writing – review & editing: EO, NK, VN, TS, RR, KPH, PK, SE.

Competing interests

The authors declare no competing interests.

Methods

Organisms Embryonic *D. melanogaster* histones, whole-genome plasmid libraries and salt gradient dialysis.

Embryonic *D. melanogaster* histone purification. The preparation of embryonic *D. melanogaster* histones octamers was carried out as described before[1,2]. In brief, 50 g of 0-12 hours old *D. melanogaster* embryos (strain OregonR) were dechorionated in 3 % sodium hypochlorite, washed with dH_20 and resuspended in 40 mL lysis-buffer (15 mM K·HEPES pH 7.5, 10 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 0.2 mM PMSF, 10 % glycerol). Embryos were homogenized (Yamamoto homogenizer), filtered through cloth and centrifuged at 6,500 *g* for 15 min. Nuclei (brownish light pellet) were washed 3 times with 50 mL sucrose-buffer (15 mM K·HEPES pH 7.5, 10 mM KCl, 5 mM $MgCl_2$, 0.05 mM EDTA, 0.25 mM EGTA, 1 mM DTT, 0.2 mM PMSF, 1.2 % sucrose) and resuspended in 30 mL sucrose-buffer containing 3 mM $CaCl_2$. To obtain mononucleosomes, nuclei were incubated for 10 min at 26 °C with 6250 Units MNase (Sigma-Aldrich). Reaction was stopped with 10 mM EDTA, nuclei were pelleted and resuspended in 6 mL TE (10 mM Tris·HCl pH 7.6, 1 mM EDTA) containing 1 mM DTT and 0.2 mM PMSF followed by 30 to 45 min of rotation at 4 °C. Nuclei were

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

octamers were eluted with 2 M KCl, concentrated and stored in 50 % glycerol and 1x Complete (Roche) protease inhibitors without EDTA at -20 °C.

Whole-genome plasmid library expansion. The *S. cerevisiae* genomic plasmid library (pGP546) was originally described by Jones et al.[3] and purchased as a clonal glycerol stock collection from Open Biosystems. Library expansion was carried out via a Singer ROTOR plating machine (Singer Instruments) (8-12 rounds, 3 replicas). After 16 hours, colonies were combined into 3x2 L of LB medium containing 50 µg/mL kanamycin and grown for 4 hours. Cells were harvested and subjected to Plasmid Giga Preparation (PC 10 000 Kit, Macherey&Nagel).

Salt gradient dialysis (SGD). For low, medium and high assembly degrees, 10 µg of plasmid library DNA (*S. cerevisiae*, *S. pombe* or *E. coli*) was mixed with ~2, 4 or 8 µg of *Drosophila* embryo histone octamers, respectively, in 100 µl assembly buffer (10 mM Tris·HCl, pH 7.6, 2 M NaCl, 1 mM EDTA, 0.05 % IGEPAL CA630, 0.2 µg BSA). For reconstitutions with precleaved DNA (Fig. 8), the plasmid library was digested with the respective restriction enzyme and purified by phenol extraction/ethanol precipitation prior to SGD. Samples were transferred to Slide-A-lyzer mini dialysis devices, which were placed in a 3 L beaker containing 300 mL of high salt buffer (10 mM Tris·HCl pH 7.6, 2 M NaCl, 1 mM EDTA, 0.05 % IGEPAL CA630, 14.3 mM β-mercaptoethanol), and dialyzed against a total of 3 L low salt buffer (10 mM Tris·HCl pH 7.6, 50 mM NaCl, 1 mM EDTA, 0.05 % IGEPAL CA630, 1.4 mM β-mercaptoethanol) added continuously via a peristaltic pump over a time course of 16 h while stirring. β-mercaptoethanol was added freshly to all buffers. After complete transfer of low salt buffer, samples were dialyzed against 1 L low salt buffer for 1 h at room temperature. DNA concentration of the SGD chromatin preparations was estimated with a DS-11+ spektrophotometer (Denovix) and could be stored at 4 °C for several weeks. To estimate the extent of the assembly degree, an aliquot of the sample was subjected to MNase digestion (as described below) for MNase-ladder read out.

Expression and purification of INO80 complex and respective mutants. Coding sequences for *S. cerevisiae* Ino80 (2xFlag), Rvb1, Rvb2, Arp5-His, Ies6 (pFBDM_1) and Actin, Arp4, Arp8, Taf14, Ies2, Ies4, Ies1, Ies3, Ies5 and Nhp10 (pFBDM_2) were subcloned into pFBDM vectors[4] and sequence verified by Sanger Sequencing (GATC Services at Eurofins Genomics). Bacmids of both vectors were generated using DH10 multibac cells[5]. Baculoviruses were generated in *Spodoptera frugiperda* (SF21) insect cells (IPLB-Sf21AE). *Trichoplusia ni* High Five (Hi5) insect cells (BTI-TN-5B1-4 Invitrogen) were co-infected with two baculoviruses 1/100 each. After 60 h cultivation at 27 °C, cells were harvested by centrifugation. For purification of the INO80 complex, cells were resuspended in lysis buffer (50 mM Tris·HCl pH 7.9, 500 mM NaCl, 10 % glycerol, 1 mM DTT, SIGMAFAST™ protease inhibitor cocktail), sonified (Branson Sonifier, 3x 20 s with 40 % duty cycle and output control 3-4) and cleared by centrifugation (Sorvall Evolution RC, SS34 rotor, 15,000 *g*). The supernatant was incubated for 1 h with anti-Flag M2 Affinity Gel (Sigma-Aldrich) and centrifuged for 15 min at 1,000 *g* and 4 °C. The anti-Flag resin was washed with buffer A (25 mM K·HEPES pH 8.0, 500 mM KCl, 10 % glycerol, 0.025 mM IGEPAL CA630, 4 mM $MgCl_2$, 1 mM DTT) and buffer B (25 mM K·HEPES pH 8.0, 200 mM KCl, 10 % glycerol, 0.02 mM IGEPAL CA630, 4 mM $MgCl_2$, 1 mM DTT). Recombinant INO80 complex was eluted with buffer B containing 1.6 mg Flag Peptide (Sigma-Aldrich). Anion exchange chromatography (MonoQ 5/50 GL, GE Healthcare, Buffer: 25 mM K·HEPES pH 8.0, 4 mM $MgCl_2$, 1 mM DTT) using a linear KCl gradient 200mM-1000mM) and, if required, size exclusion chromatography (Superose 6, 10/300 GL, 25 mM K·HEPES pH 8.0, 200mM, 4 mM $MgCl_2$, 1 mM DTT) was used for further purification which resulted in a monodisperse INO80 complex (Figure S1A,B,E). Using standard cloning techniques, three INO80 (2xFlag) HSA domain mutants (HQ1, HQ2, HQ1/2; Figure 2C, S1E), one N-terminal deletion mutant ($Ino80^{\Delta N}$, deletion of the first 461 amino acids of the N-terminus of Ino80) and two INO80 (2xFlag) Nhp10 module mutants (ΔNhp10 (INO80 complex without Ies1, Ies3, Ies5 and Nhp10 but with Ino80 N-terminus) and HMGII (Figure 2C, S1E) pFBDM vectors were generated and integrated into baculoviruses using MultiBac Technology as described above. Expression and purification of mutant INO80 complexes was essentially carried out as WT INO80 complex purification. The INO80 core complex from *Chaetomium thermophilum* (equivalent to the *S. cerevisiae* N-terminal deletion mutant) was essentially purified as described in [6].

Genome-wide remodeling reaction. All remodeling reactions were performed at 30 °C in 100 µL with final buffer conditions of 26.6 mM Na·HEPES pH 7.5, 1 mM Tris·HCl pH 7.6, 85.5 mM NaCl, 8 mM KCl, 10 mM ammonium sulfate, 10 mM creatine phosphate (Sigma-Aldrich), 3 mM $MgCl_2$, 2.5 mM ATP, 0.1 mM EDTA, 0.6 mM EGTA, 1 mM DTT, 14 % glycerol, 20 ng/µl creatine kinase (Roche Applied Science). Remodeling reactions were started by adding 10 µL SGD chromatin corresponding to ~ 1 µg DNA assembled into nucleosomes and terminated by adding 0.8 Units apyrase (NEB) followed by incubation at 30 °C for 30 min. Independent replicates of remodeling reactions refer to independent SGD chromatin preparations. The experimental conditions for each sample are detailed in Supplementary Data 1.

MNase-seq. After apyrase addition, remodeling reactions were supplemented with $CaCl_2$ to a final concentration of 1.5 mM and digested with 100 Units MNase (Sigma) to generate mostly monoucleosomal DNA. 10 mM EDTA and 0.5 % SDS (final concentrations) were added to stop the MNase digest. After proteinase K treatment for 30 min at 37 °C, samples were ethanol precipitated and electrophoresed for 1.5 - 2 h at 100 V using a 1.5 % agarose gel in 1x Tris-acetate-EDTA (TAE) buffer. Mononucleosome bands were excised and purified with PureLink Quick Gel Extraction Kit (ThermoFisher Scientific).
For library preparation, 10-50 ng of mononucleosomal DNA was incubated with 1.25 Units Taq polymerase (NEB), 3.75 Units T4 DNA polymerase (NEB) and 12.5 Units T4-PNK (NEB) in 1x ligation buffer (B0202S, NEB) for 15 min at 12 °C, 15 min at 37 °C and 20 min at 72 °C. To ligate NEBNext Adaptors (0.75 µM final concentration, NEBNext Multiplex Oligos Kit) to the DNA, samples were incubated with T4 DNA ligase (NEB) at 25 °C for 15 min, followed by incubation with 2 Units USER enzyme (NEB) for 10 min at 37 °C. Fragments were purified using 2 volumes AMPure XP beads (Beckman

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Coulter) and amplified for 8-10 cycles using NEBNext Multiplex Oligos, Phusion High-Fidelity DNA Polymerase (1 U, NEB), deoxynucleotide solution mix (dNTP, 2.5 mM, NEB) and Phusion HF Buffer (1x, NEB). The following protocol was applied for amplification: 98 °C for 30 s, 98 °C for 10 s, 65 °C for 30 s, 72 °C for 30 s with a final amplification step at 72 °C for 5 min. DNA content was assessed by using Qubit dsDNA HS Assay Kit (Invitrogen). PCR reactions were applied to an 1.5 % agarose gel, needed fragment length (~270 bp) was excised and purified via PureLink Quick Gel Extraction Kit (ThermoFisher Scientific). DNA was measured again with Qubit dsDNA HS Assay Kit and diluted to a final concentration of 10 nM (calculation based on the assumption that the DNA fragment length is 272 bp, i.e., 147 bp nucleosomal DNA and 122 bp sequencing adaptor). Diluted samples were pooled according to sequencing reads (~6 Mio reads/ sample). The final pool was quantified with BioAnalyzer (Agilent) and analyzed on an Illumina HiSeq 1500 in 50 bp single-end mode (Laboratory for Functional Genome Analysis, LAFUGA, LMU Munich).

Expression and purification of human tailless histone octamers. The genes for expression of tailless human histones H2A, H2B and H4 were cloned in pET21b vectors (Merck, Darmstadt, Germany) by blunt-end ligation of genes coding for full-length human histones. The gene coding for human tailless H3 was cloned in a pETM-11 vector (kindly provided by EMBL, Heidelberg, Germany) carrying a N-terminal SUMO-tag by Gibson assembly[7]. The SUMO-tag was removed during octamer assembly. Constructs of tailless histones were designed according to globular domains identified by tryptic digest of full-length histone[8-10] and comprised the following amino acids: H2A: 13 – 118; H2B: 24 – 125; H3: 27 – 135; H4: 20 – 102. Histones were purified by a combination of inclusion body purification and ion-exchange chromatography, essentially as described previously[11,12]. In brief, histones were expressed in *E. coli* BL21 (DE3) cells (Merck, Darmstadt, Germany) for 2 h after induction with 1 mM IPTG at 37 °C and disrupted under non-denaturing conditions to separate inclusion bodies from lysate. Inclusion bodies were first washed with 1% Triton-X100. Subsequently, inclusion bodies were resuspended in 7 M guanidinium chloride and dialyzed against 8 M urea. Individual histones were purified by cation-exchange chromatography, refolded under low-salt conditions and polished by anion-exchange chromatography. For long-time storage, histones were lyophilized overnight. For octamer reconstitution, histones were resuspended in 25 mM Tris, pH 7.5, 7 M guanidinium chloride, 0.25 mM DTT, mixed at 1.2-fold excess of H2A and H2B and dialyzed against 25 mM Tris·HCl pH 7.5, 2 M NaCl, 0.25 mM DTT overnight. 1 mg/mL SENP2 protease was added after 3 h. The octamer of tailless histones was purified by size-exclusion chromatography using a Superdex 200 16/60 column (GE Healthcare), which separated the octamer from aggregate, H2A/H2B dimers, the SENP2 protease and the SUMO-tag. The purification was analyzed on a 18 % polyacrylamide SDS gel stained with Coomassie (data not shown). The octamer was concentrated to 3.0 mg/mL and stored at -20°C in 50% glycerol.

Expression and purification of *S. cerevisiae* Reb1. For genome-wide remodeling reaction *S. cerevisiae* Reb1 was purified exactly as described in[13]. For ATPase and mononucleosome sliding assays Reb1 was purified as follows: Reb1 was amplified from BY4741 genomic *S. cerevisiae* DNA by PCR and cloned into pET21b (Novagen) via InFusion cloning (Clontech) with a Streptavidin tag at the C terminus. Correct sequences were verified via Sanger sequencing (GATC Services at Eurofins Genomics). Expression plasmids were transformed into BL21 (DE3) cd^+ cells. Three liters of LB medium supplemented with 600 mg/L ampicillin were inoculated with 200 mL pre-culture. Cells were grown at 37 °C to an OD_{600} of 0.6 (WPA CO8000 cell density meter). Induction was carried out by addition of IPTG to a final concentration of 1 mM. Cells were grown overnight at 18 °C, harvested by centrifugation (3,500 rpm, Sorvall Evolution RC) and stored at -80 °C. Cells were resuspended in lysis buffer (50 mM Tris·HCl pH 7.9, 500 mM NaCl, 7 % glycerol, 1 mM DTT, 7 % sucrose and protease inhibitor 1:100), sonicated (Branson Sonifier 250, 5 min at 40-50 % duty cycle and output control 4) and cleared by centrifugation (Sorvall Evolution RC, SS34 rotor, 15,000 *g*). The supernatant was dialyzed over night against 2 L low salt buffer (25 mM K·HEPES pH 8.0, 50 mM KCl, 7 % glycerol, 4 mM $MgCl_2$, 1 mM DTT). Heparin chromatography (5 mL column, elution buffer: 25 mM K·HEPES pH 8.0, 1 M KCl, 7 % glycerol, 4 mM $MgCl_2$, 1 mM DTT) followed by size exclusion chromatography (Superdex 200 10/300, buffer: 25 mM K·HEPES pH 8.0, 200 mM KCl, 7 % glycerol, 4 mM $MgCl_2$, 1 mM DTT) were used for purification. Peak fractions were analyzed by Coomassie SDS-PAGE. Fractions containing Reb1 were pooled, concentrated and stored at -80 °C.

Preparation of mononucleosomes with recombinant human octamers. Canonical human histones were provided by The Histone Source – Protein Expression and Purification (PEP) Facility at Colorado State University. Lyophilized individual human histones were resuspended in 7 M guanidinium chloride, mixed at a 1.2-fold molar excess of H2A/H2B and dialyzed against 2 M NaCl for 16 h. Histone octamers were purified by size exclusion chromatography (HILoad 16/600 Superdex 200 column, GE Healthcare) and stored at -20 °C in 50 % glycerol.
We used fluorescein-labeled Widom 601 DNA[14] with 80 bp extranucleosomal DNA (0N80 orientation) harboring an in vivo ChIP-Exo verified Reb1 binding site[15] of *S. cerevisiae* gene yGL167c (Reb1 binding motif: TTACCC) 64 or 84 bp distant to the 601 sequence. The DNA template (yGL267c_601) was amplified via PCR, purified by anion exchange chromatography (HiTrap DEAE FF, GE Healthcare) and vacuum concentrated. DNA and assembled histone octamer were mixed in 1.1-fold molar excess of DNA at 2 M NaCl. Over a time-period of 17 h at 4 °C the NaCl concentration was reduced to a final concentration of 50 mM NaCl. Again, anion exchange chromatography was used to purify reconstituted nucleosome core particle (NCP) which were then dialyzed to 50 mM NaCl. NCPs were concentrated to 1 mg/mL and stored at 4 °C.

ATPase Assay. As described previously[16], we applied an NADH-based ATPase assay [17] to determine INO80's ATPase rate. 15 nM INO80 were incubated at 30 °C in a final volume of 50 µl assay buffer (25 mM K·HEPES pH 8.0, 50 mM KCl, 5 mM $MgCl_2$, 0.1 mg/mL BSA) with 0.5 mM phosphoenolpyruvate, 2 mM ATP, 0.2 mM NADH and 25 units/mL lactate dehydrogenase/pyruvate kinase (Sigma-Aldrich) to monitor the NADH dependent fluorescence signal in non-binding, black, 384-well plates (Greiner) at an

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

excitation wavelength of 340 nm and an emission wavelength of 460 nm over a 40-min period. We used the Tecan Infinite M1000 (Tecan) plate reader for read out. For all samples, ATPase activity was determined at maximum INO80 WT ATPase activity. ATPase activity was stimulated with 25 nM GL167c-0N80 mononucleosomes with or without equimolar ratios WT Reb1. Using maximal initial linear rates corrected for the buffer blank, we calculated final ATP turnover rates.

Mononucleosome sliding assay. Nucleosome sliding activity of INO80 wild type and mutant complexes were monitored on Reb1 site-0N80 mononucleosomes in absence and presence of Reb1. INO80 at a concentration of 10 nM was incubated with 90 nM of Reb1 site-0N80 mononucleosomes in sliding buffer at 26 °C (sliding buffer: 25 mM Na·HEPES pH 8.0, 60 mM KCl, 7 % glycerol, 0.10 mg/mL BSA, 0.25 mM dithiothreitol and 2 mM $MgCl_2$). ATP and $MgCl_2$ at final concentrations of 1 mM and 2 mM, respectively, were added to start the sliding reaction. After 30 s, 60 s, 120 s, 300 s, 600 s, 1800 s and 3600 s the reaction was stopped by adding lambda DNA (NEB) to a final concentration of 0.2 mg/mL. To separate distinct nucleosome species, we applied NativePAGE (NativePAGE Novex 4-16 % Bis-Tris Protein Gels, Invitrogen). The fluorescein-labeled mononucleosomal DNA was visualized by a Typhoon™ FLA 9000 imager.

Data Processing. Sequencing data was mapped to the SacCer3 (R64) genome using *bowtie*[18]. Multiple matches were omitted. After mapping, data was imported into R Studio using GenomicAlignments[19]. Every read was shifted by 73 bp to cover the nucleosome dyad and extended to 50 bp. Genome coverage was calculated, and aligned to either *in vivo* +1 nucleosome positions[20], BamHI cut sites, Reb1 SLIM-ChIP hits[21] or Reb1 PWM hits[22]. Signal was normalized per gene in a 2001 bp window centered on the alignment point.

Heatmaps were sorted either by NFR length (distance between in vivo +1 and -1 nucleosome annotated by calling nucleosomes of *in vivo* MNase-seq data, see below) or by Reb1 binding score. For the latter, Reb1 SLIM-ChIP data (GSM2916407) was aligned to in vivo +1 nucleosome positions and sorted by signal strength in a 120 bp-window 160 bp upstream of every +1 nucleosome.

For promotor grouping according to Reb1 site orientation, Reb1 SLIM-ChIP hits which contain a PWM site (± 50 bp) and which are located within 400 bp upstream of in vivo +1 nucleosomes were used. Cluster 1 contains promotors where the Reb1 PWM motif is located on the sense strand and cluster 2, where the Reb1 PWM motif is located on the antisense strand. Cluster 3 contains Reb1 sites at bidirectional promotors.

DNA shape and poly(dA:dT) analysis surrounding Reb1 binding sites. The DNA sequence of the yeast genome (SacCer3) was downloaded from *Saccharomyces* Genome Database (SGD) and the DNA shape feature scores (helix twist, propeller twist, minor groove width and electrostatic potential) were calculated for the entire genome using the R package *DNAshapeR* (v1.10.0). Similar to[13], the resulting DNA shape vectors were smoothed with a 5-bp rollmean. For composite analysis, DNA shape feature specific values were extracted in a window of -2000 to 2000 bp around Reb1 binding sites, oriented with respect to Reb1 motif directionality, and averaged by base pair. Plotted distance around Reb1 features are indicated in respective figures.

For the poly(dA:dT) analysis, stretches of 6 nucleotide long polyA (5'-AAAAAA-3') or polyT (5'-TTTTTT-3') were identified in the yeast genome using R package *Biostrings* (v2.52.0) and counted. For composite analysis, ploy(dA) or poly(dT) counts were extracted in a window of -2000 to 2000 bp around Reb1 binding sites, oriented with respect to Reb1 motif directionality, and averaged by base pair. Plotted distance around Reb1 features are indicated in respective figures.

Identification of TSS +1 nucleosomes. +1 nucleosome positions were called according to [23]. In more detail, mononucleosomal fragments generated from BY4741 MNase digested chromatin were sequenced on an Illumina Genome analyzer, mapped to the SacCer3 genome with *bowtie* [18] and shifted by 73 bp with respect to sequencing read directionality to obtain theoretical nucleosome dyads. The obtained dyad-density counts were smoothed with sliding Gaussian filter (width = 100, mean = 0, SD = 25) and resulting values were sorted by decreasing values. Iteratively, the position with the highest value was added to the list of "dyad centers" and all values for positions within +/-120 bp surrounding the position with the highest value were removed from further analysis. The top 90% of nucleosome dyad centers, by value, constituted the final list of nucleosome positions. Plus 1 nucleosome dyad positions were defined as the nearest nucleosome dyad position to TSS within a window 0 to +500 bp from the TSS, with respect to direction of transcription.

Genome-wide principal component and DNA shape analysis of nucleosomes. For PCA and DNA shape analysis, mononucleosomes were sequenced in 50 bp paired-end mode on an Illumina HiSeq1500. If not stated otherwise, functions were called with default parameters. Read pairs were aligned using bowtie2 (version 2.2.9) with options "-X 250 --no-discordant --no-mixed --no-unal". Only unique matches were kept, and orphaned mates removed. Nucleosomes were called on each sample using bioconductor/nucleR (2.16.0) on nucleosomal fragments defined by paired reads as follows: fragments were processed with trimming to 40 bp around the dyads and their coverage was calculated. Noise was removed using FFT filtering with parameter pcKeepComp=0.02 and peak detection was carried out with threshold 99%.

For each sample in an analysis set, sample-specific dyad positions obtained by nucleosome calling were enlarged to 20 bp and all positions were merged across the samples. Overlapping regions were joined. We excluded regions locating closer than 250 bp to tile borders and those residing in a region with high artifactual signals (chr III, 91000-93000 bp).

On this joint set of nucleosome dyads, we counted the number of overlapping fragments (reduced to their center position) for each sample. With x being the number of counts of sample-specific fragment centers overlapping one dyad region of the joint set and sum(x) being the sum of all counts across all dyad regions in the sample the data was normalized using the formula: normalized occupancy (dyad region) = log2(((x/sum(x))*1000)+0.001). The resulting matrix was subjected to principal component analysis. K-means clustering was applied to the resulting principal

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

components to group nucleosomes based on similar occupancy patterns across sample conditions.

DNA shape features in windows of 320 bp around dyad positions were calculated with bioconductor/DNAshapeR (version 1.14.0). DNA rigidity scores of each position in windows of 320 bp around dyad positions were calculated as the length of the longest consecutive A_nT_m (n≥0, m≥0 and $n+m$≥2) sequence element that contains this position.

Data Availability. All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus under accession numbers GSE145093 and GSE140614.

References for Methods

1. Krietenstein, N., Wippo, C.J., Lieleg, C. & Korber, P. Genome-wide in vitro reconstitution of yeast chromatin with in vivo-like nucleosome positioning. *Methods Enzymol* **513**, 205-32 (2012).

2. Simon, R.H. & Felsenfeld, G. A new procedure for purifying histone pairs H2A + H2B and H3 + H4 from chromatin using hydroxylapatite. *Nucleic Acids Res* **6**, 689-96 (1979).

3. Jones, G.M. et al. A systematic library for comprehensive overexpression screens in Saccharomyces cerevisiae. *Nature Methods* **5**, 239-241 (2008).

4. Trowitzsch, S., Bieniossek, C., Nie, Y., Garzoni, F. & Berger, I. New baculovirus expression tools for recombinant protein complex production. *J Struct Biol* **172**, 45-54 (2010).

5. Berger, I., Fitzgerald, D.J. & Richmond, T.J. Baculovirus expression system for heterologous multiprotein complexes. *Nat Biotechnol* **22**, 1583-7 (2004).

6. Eustermann, S. et al. Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* **556**, 386-390 (2018).

7. Gibson, D.G. Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. *Nucleic Acids Res* **37**, 6984-90 (2009).

8. Bohm, L., Crane-Robinson, C. & Sautiere, P. Proteolytic digestion studies of chromatin core-histone structure. Identification of a limit peptide of histone H2A. *Eur J Biochem* **106**, 525-30 (1980).

9. Bohm, L., Briand, G., Sautiere, P. & Crane-Robinson, C. Proteolytic digestion studies of chromatin core-histone structure. Identification of the limit peptides of histones H3 and H4. *Eur J Biochem* **119**, 67-74 (1981).

10. Bohm, L., Briand, G., Sautiere, P. & Crane-Robinson, C. Proteolytic digestion studies of chromatin core-histone structure. Identification of limit peptides from histone H2B. *Eur J Biochem* **123**, 299-303 (1982).

11. Dyer, P.N. et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol* **375**, 23-44 (2004).

12. Klinker, H., Haas, C., Harrer, N., Becker, P.B. & Mueller-Planitz, F. Rapid purification of recombinant histones. *PLoS One* **9**, e104029 (2014).

13. Krietenstein, N. et al. Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell* **167**, 709-721 e12 (2016).

14. 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).

15. Rhee, H.S. & Pugh, B.F. ChIP-exo method for identifying genomic location of DNA-binding proteins with near-single-nucleotide accuracy. *Curr Protoc Mol Biol* **Chapter 21**, Unit 21 24 (2012).

16. Knoll, K.R. et al. The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling. *Nat Struct Mol Biol* **25**, 823-832 (2018).

17. Kiianitsa, K., Solinger, J.A. & Heyer, W.D. NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Anal Biochem* **321**, 266-71 (2003).

18. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).

19. Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLoS Comput Biol* **9**, e1003118 (2013).

20. Xu, Z. et al. Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**, 1033-7 (2009).

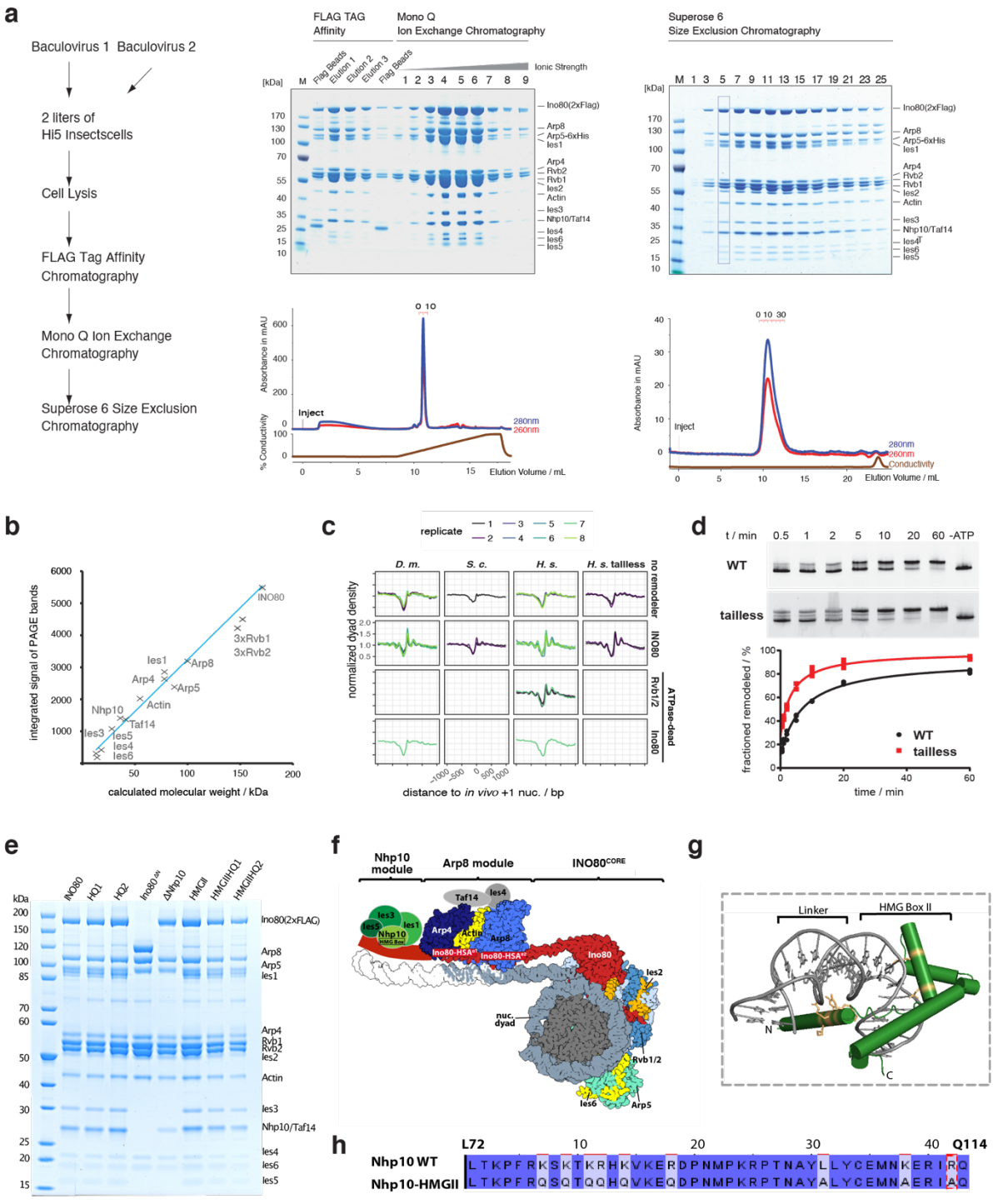
21. Gutin, J. et al. Fine-Resolution Mapping of TF Binding and Chromatin Interactions. *Cell Rep* **22**, 2797-2807 (2018).

22. Badis, G. et al. A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. *Mol Cell* **32**, 878-87 (2008).

23. Tirosh, I. Computational analysis of nucleosome positioning. *Methods Mol Biol* **833**, 443-9 (2012).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

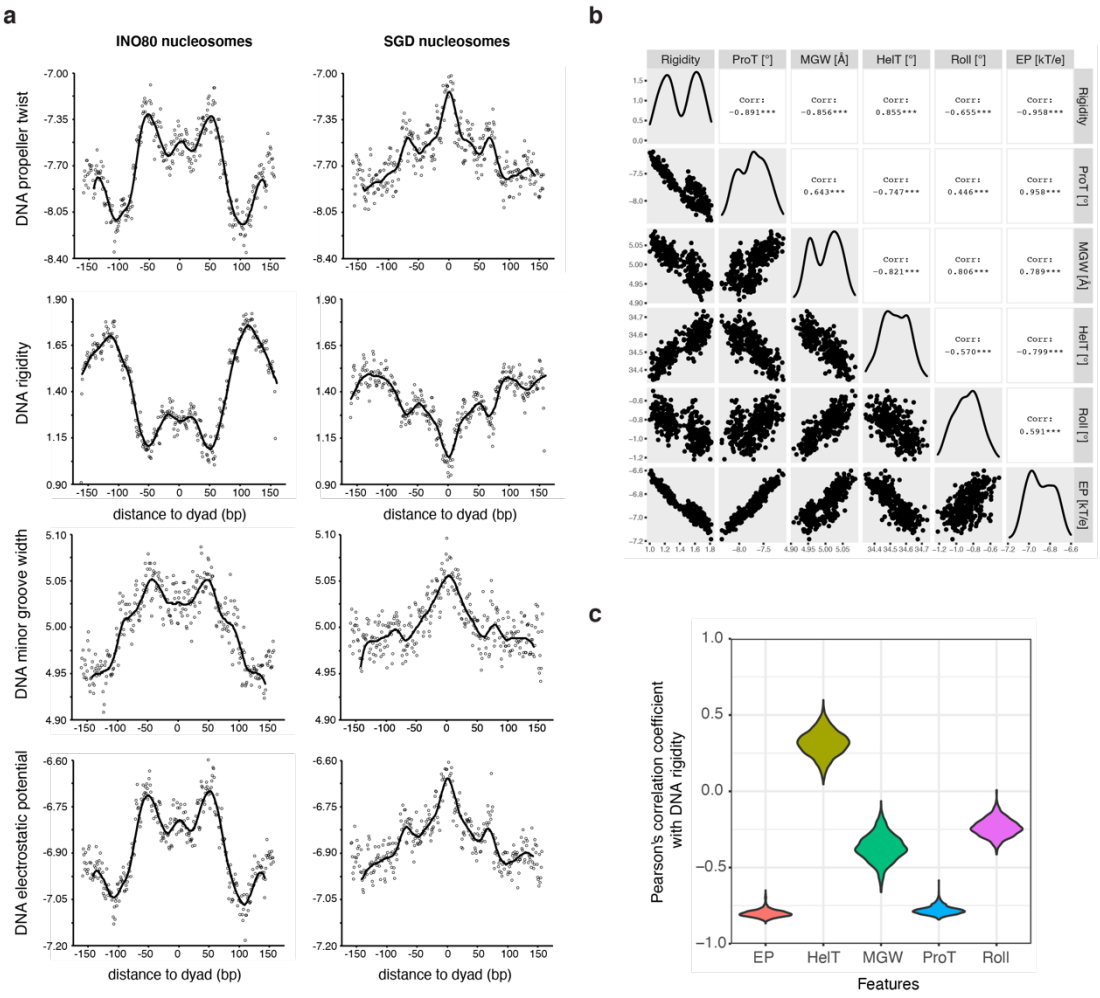
Supplementary Figures



Supplementary Figure 1. Expression, purification and activity of recombinant INO80. a Recombinant expression and purification of 15-subunit *S. cerevisiae* INO80 complex. Left: Schematic of expression and purification work flow. Two baculoviruses encoding five (Ino80, Arp5, Ies6, Rvb1 and Rvb2) and ten INO80 subunits (Ies1-5, Nhp10, Taf14, Actin, Arp4, Arp8), respectively, were used for insect cell expression. Middle and right: SDS-PAGE analysis of indicated chromatographies. Numbered lanes indicate elution fractions matching chromatograms below gels. Boxed lane represents a fraction used in this study. **b** Quantification of Coomassie-stained SDS PAGE bands shows stoichiometric assembly of recombinant *S.cerevisiae* INO80 complex. Note that AAA+ ATPase Rvb1 and Rvb2 form a hetero-hexamer. **c** Composite plots of MNase-seq data of individual replicates for the indicated combinations of histones (columns) and remodeling enzymes (rows). **d** top: Native gel electrophoresis analysis at indicated time points of mononucleosome sliding assay kinetics with wild type (WT) or

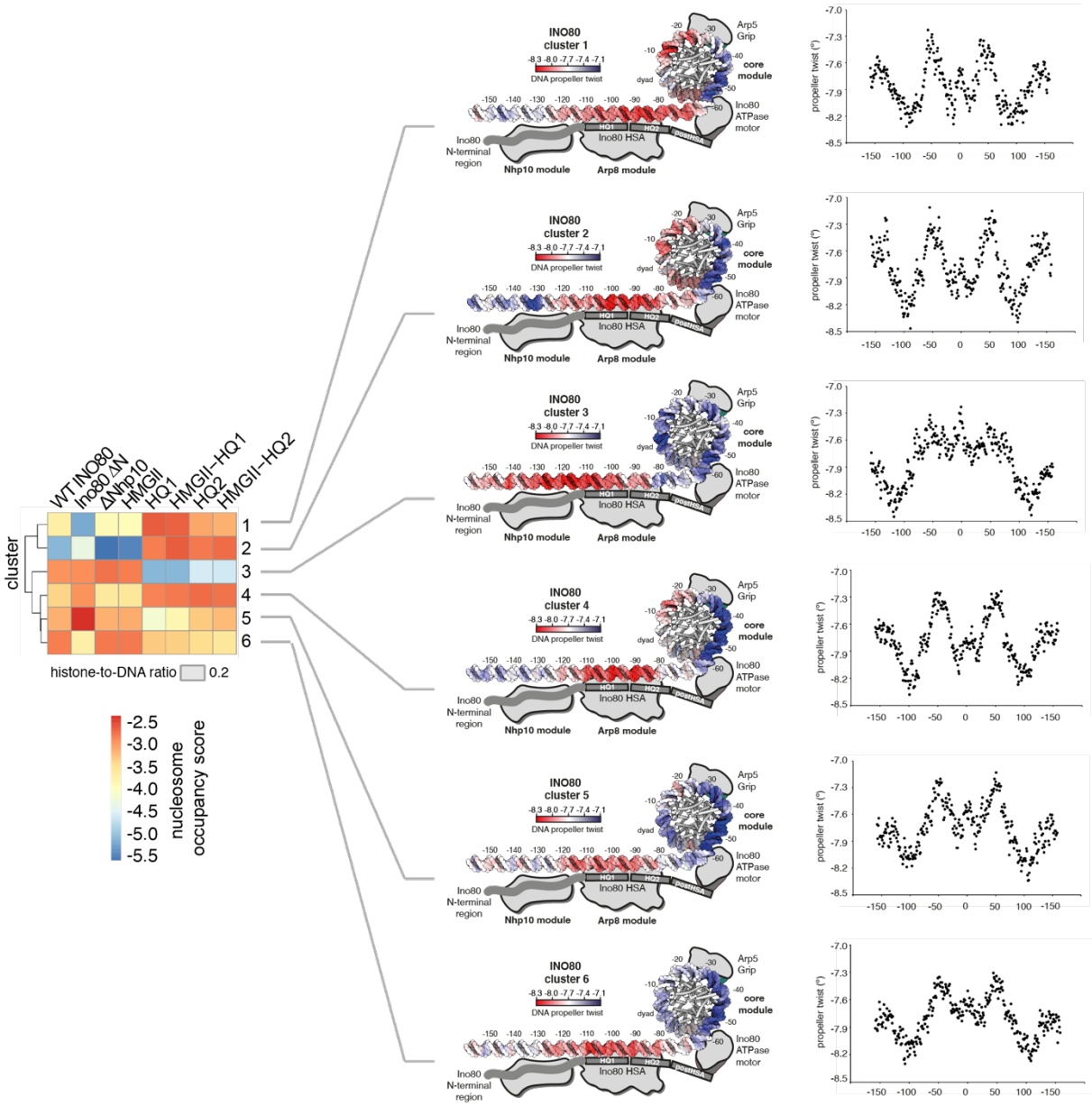
bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tailless (tailless) recombinant *H. sapiens* histones and wild type recombinant *S. cerevisiae* INO80 complex. "-ATP" denotes 60 min time point without ATP. bottom: Quantification of data from top. **e** SDS-PAGE analysis of purified, recombinant WT (INO80) or indicated mutant complexes. **f** left: Structure-based [6,16] model of a nucleosome bound by the INO80 complex with indicated subunits. Taf14, Ies4 and Nhp10 module organization is assumed. **g** Model of Nhp10 HMG box-like and Linker region (residues 62-172) based on TFAM structure (pdb 3tq6). **h** Sequence alignment showing mutated residues in Nhp10-HMGII mutant. Panels e-h are also shown in the accompanying paper Oberbeckmann & Niebauer et al.



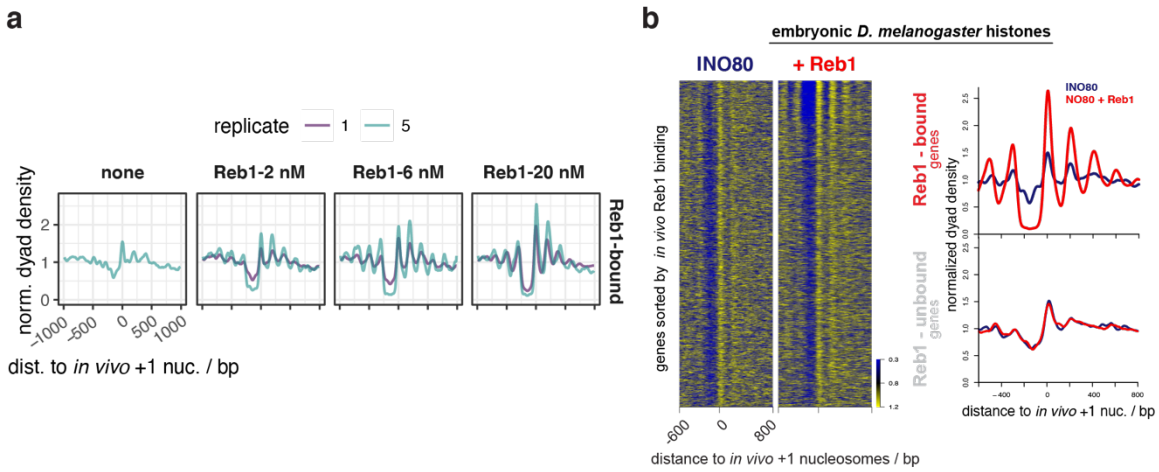
Supplementary Figure 2. DNA shape/mechanics features of INO80 and SGD positioned nucleosomes. a DNA shape/mechanics profiles (DNA propeller twist, DNA rigidity, DNA minor groove width and DNA electrostatic potential) derived from INO80 and SGD positioned nucleosomes. **b** Pearson's correlation coefficients between six DNA features: minor groove width (MGW), helix twist (HelT), propeller twist (ProT), Roll, Electrostatic potential (EP), and DNA rigidity. The average profiles of DNA features across all nucleosomal sequences are used to obtain the correlation coefficients between features. **c** Violin plot of Pearson's correlation coefficients between DNA rigidity and other DNA features of all nucleosomal sequences. The coefficient is obtained by correlating the DNA feature profiles of each sequence individually.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplementary Figure 3. Clustering and DNA shape analysis of WT and mutant INO80. DNA propeller twist shape profile of nucleosomal DNA sequences. Color-coded mapping is shown for each cluster.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.03.366690; this version posted November 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplementary Figure 4. Nucleosome positioning in presence of Reb1. a Composite plots of MNase-seq data for individual replicates of samples as in Figure 6a,b, but only for genes with promoter Reb1 sites (Reb1-bound, same as red shading in Figure 6a) and also including SGD chromatin incubated with INO80 in the absence of Reb1 (none). **b** As Figure 6a,b but for the SGD chromatin with embryonic *D. melanogaster* histones at histone-to-DNA mass ratio of 0.4 and only 20 nM Reb1 (+ Reb1).

CHAPTER 4: RULER ELEMENTS IN CHROMATIN REMODELERS SET NUCLEOSOME ARRAY SPACING AND PHASING

Elisa Oberbeckmann[1,5], Vanessa Niebauer[2,5], Shinya Watanabe[3], Lucas Farnung[4], Manuela Moldt[2], Andrea Schmid[1], Patrick Cramer[4], Craig L. Peterson[3], Sebastian Eustermann[2], Karl-Peter Hopfner[2] and Philipp Korber[1]

[1]Division of Molecular Biology, Biomedical Center, Faculty of Medicine, Ludwig-Maxmilians-Universität München, Martinsried near to Munich, Germany; [2]Gene Center, Faculty of Chemistry and Pharmacy, Ludwig-Maxmilians-Universität München, Munich, Germany; [3]Program of Molecular Medicine, University of Massachusetts, Worcester, USA; [4]Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany; *[5]These authors contributed equally to the work.*

This chapter is not published in a peer-reviewed journal yet
but under revision and made public on BioRxiv.

Author contributions to "Ruler elements in chromatin remodelers set nucleosome array spacing and phasing"

Conceptualization: P.K., S.E., E.O., K.-P.H.; Data curation: E.O.; Formal analysis: E.O.; Funding acquisition, Project administration, Supervision: P.K., K.-P.H., S.E., P.C., C.L.P.; Investigation: E.O., V.N., S.W., L.F., M.M., A.S.; Methodology: E.O., V.N., S.E., S.W., L.F., P.K.; Validation: E.O., V.N., M.M., A.S., S.W., L.F., P.K., S.E.; Visualization: E.O., V.N., P.K., S.W., L.F.; Writing original draft: P.K., S.E., E.O.; Writing – review & editing: P.K., S.E., E.O., V.N., K.-P.H., L.F., S.W., C.L.P., P.C.

---

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

Ruler elements in chromatin remodelers set nucleosome array spacing and phasing

Elisa Oberbeckmann[1,7], Vanessa Niebauer[2,7], Shinya Watanabe[3], Lucas Farnung[4], Manuela Moldt[2], Andrea Schmid[1], Patrick Cramer[4,6], Craig L. Peterson[3,6], Sebastian Eustermann[2,5,6]*, Karl-Peter Hopfner[2,6]*, Philipp Korber[1,6,8]*

[1]Division of Molecular Biology, Biomedical Center, Faculty of Medicine, Ludwig-Maxmilians-Universität München, Martinsried near to Munich, Germany; [2]Gene Center, Faculty of Chemistry and Pharmacy, Ludwig-Maxmilians-Universität München, Munich, Germany; [3]Program of Molecular Medicine, University of Massachusetts, Worcester, USA; [4]Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany; [5]current address: European Molecular Biology Laboratory (EMBL), Structural and Computational Biology Unit, Heidelberg, Germany

Arrays of regularly spaced nucleosomes dominate chromatin and are often phased by alignment to reference sites like active promoters. How the distances between nucleosomes (spacing), and between phasing sites and nucleosomes are determined remains unclear, and specifically, how ATP dependent chromatin remodelers impact these features. Here, we used genome-wide reconstitution to probe how *Saccharomyces cerevisiae* ATP dependent remodelers generate phased arrays of regularly spaced nucleosomes. We find that remodelers bear a functional element named the 'ruler' that determines spacing and phasing in a remodeler-specific way. We use structure-based mutagenesis to identify and tune the ruler element residing in the Nhp10 and Arp8 modules of the INO80 remodeler complex. Generally, we propose that a remodeler ruler regulates nucleosome sliding direction bias in response to (epi)genetic information. This finally conceptualizes how remodeler-mediated nucleosome dynamics determine stable steady-state nucleosome positioning relative to other nucleosomes, DNA bound factors, DNA ends and DNA sequence elements.

Nuclear DNA is packaged into chromatin based on a repeating building block, the nucleosome core particle (NCP; (Kornberg, 1974; Olins and Olins, 1974)), where 147 base pairs (bp) of DNA are wound around a histone protein octamer (Kornberg and Lorch, 1999; Luger et al., 1997; Olins and Olins, 2003). Packaging by nucleosomes orchestrates all genomic processes (Lai and Pugh, 2017).

Nucleosomes mainly occur in regular arrays where they are aligned to each other such that the lengths of linker DNA between NCPs are about constant within an array. Linker lengths may vary among arrays in the same cell (Baldi et al., 2018b; Chereji et al., 2018; Ocampo et al., 2016; Valouev et al., 2011) and differ on average between cell types and species (van Holde, 1989). Arrays are often phased, i.e., aligned relative to a genomic reference point. A combination of both *in vivo* studies (Ganapathi et al., 2011; Hartley and Madhani, 2009; Kubik et al., 2018; Tsankov et al., 2011; van Bakel et al., 2013; Yan et al., 2018; Yarragudi et al., 2004) and *in vitro* reconstitutions (Krietenstein et al., 2016) indicated that these genomic alignment points or "barriers" often reflect the binding of abundant, sequence-specific DNA binding proteins, like the general regulatory factor (GRFs) Reb1, Abf1, or Rap1 in budding yeast or other architectural factors like CTCF in mammals (Wiechens et al., 2016) or Phaser in flies (Baldi et al., 2018a).

Throughout eukaryotes, phased arrays are prominent at active promoters. Nucleosome-depleted regions (NDRs) at the core promoter are flanked by arrays that begin with the so called +1 nucleosome close to the transcription start site (TSS) and cover the gene body (Baldi et al., 2020; Lai and Pugh, 2017). This organization is important for transcription fidelity as mutants with impaired array phasing show aberrant transcription initiation (Challal et al., 2018; Hennig et al., 2012; Kubik et al., 2019; Pointner et al., 2012; Smolle et al., 2012). While nucleosome arrays are likely the most pervasive and longest known chromatin organization, their generation is still not explained. Specifically, regular spacing requires fixed distances between

[6] Senior author
[7] These authors contributed equally.
[8] Lead contact
* Correspondence: pkorber@lmu.de, hopfner@genzentrum.lmu.de, sebastian.eustermann@embl.de

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

nucleosomes, and phasing requires a fixed distance between array and reference point. What sets these distances?

In vivo and *in vitro* data suggest that ATP dependent chromatin remodeling enzymes (remodelers) are key to the answer. Remodelers are conserved in eukaryotes (Flaus et al., 2006) and mobilize, reconfigure, or disassemble/reassemble nucleosomes upon ATP hydrolysis (Clapier and Cairns, 2009; Clapier et al., 2017). They are subdivided into the SWI/SNF, ISWI, CHD, and INO80 families, according to their main ATPase sequence features. Besides the core ATPase, remodelers often contain additional domains and subunits that bind the nucleosome, regulate activity and targeting, and convert their DNA tracking activity into the remodeler-specific chemo-mechanical reaction. For example, nucleosome disassembly is accomplished only by SWI/SNF family members and histone exchange only by INO80 family members, while nucleosome sliding is catalyzed by most remodelers.

Particularly relevant for array generation is an ATP-dependent nucleosome spacing activity, by which some remodelers convert irregular arrays into arrays of regularly spaced nucleosomes. Remodelers of the ISWI, CHD, and INO80 (Ito et al., 1997; Tsukiyama et al., 1999; Udugama et al., 2011; Varga-Weisz et al., 1997), but not of the SWI/SNF family, show spacing activity. This activity was suggested to rely on a length-sensor mechanism (Yang et al., 2006; Zhou et al., 2018) where nucleosome sliding rate is regulated by linker DNA length. Sliding one nucleosome back and forth between two other nucleosomes, with a linker length-dependent velocity, would center a nucleosome at steady state when both flanking linkers have the same length.

While the length-sensor mechanism may equalize linker lengths and thereby generate spacing distance *regularity*, it does not by itself determine spacing distance *length* in absolute terms. This would reciprocally depend on nucleosome density. However, spacing *in vivo* (Gossett and Lieb, 2012; Hennig et al., 2012; van Bakel et al., 2013), as well as generated *in vitro* (Lieleg et al., 2015; Zhang et al., 2011), remained constant despite changes in nucleosome density. This was called "active packing" (Zhang et al., 2011) or "clamping" (Lieleg et al., 2015), but it remained unclear if remodeler or nucleosome features led to such density-independent spacing.

Structural studies suggested that the yeast ISW1a remodeler contacts a neighboring nucleosome and may set the linker length by a "protein ruler" (Yamada et al., 2011). Two ISWI family remodelers, yeast ISW1a and ISW2, each generated regular arrays aligned at DNA-bound Reb1 or Abf1 *in vitro*, but with different spacing at the same nucleosome density (Krietenstein et al., 2016). This points towards a remodeler-specific linker length determining ruler mechanism. Also suggestive of a built-in ruler, INO80 required a minimum linker length for nucleosome sliding (Zhou et al., 2018) and recognized linker DNA via a structural module that was important for sliding (Knoll et al., 2018).

The ruler metaphor may indeed describe a remodeler mechanism that measures and sets the phasing and spacing distances of arrays. However, so far it is mainly suggestive and has to be substantiated in molecular terms. This would be exceedingly convoluted *in vivo* but requires a defined system that allows to assay the generation of phased regular arrays by remodelers and to dissect if and how a ruler mechanism is at work. Are there rulers within some or all remodelers with spacing activity? Are linker length vs. distance to barrier determined in the same or different way? Are rulers autonomous or does the outcome depend on nucleosome density or underlying DNA sequence? Ultimately, is it possible to tune a ruler, i.e., can a remodeler be mutated to generate arrays with altered spacing and/or phasing distances?

Here, we used genome-wide *in vitro* chromatin reconstitution with purified remodelers ((Krietenstein et al., 2016), accompanying paper Oberbeckmann & Krietenstein et al.) to answer these questions. All yeast remodelers with spacing activity, ISW1a, ISW2, Chd1, and INO80 have rulers that are largely autonomous regarding underlying DNA sequence but some may respond to nucleosome density. Remodeler-specific rulers mechanistically explain earlier *in vivo* observations. Structure-guided mutations in recombinant INO80 complexes led to shorter or longer spacing and phasing distances and showed that these quantities may be uncoupled. Finally, we propose a model how remodeler rulers position nucleosomes by regulating sliding direction bias according to (epi)genetic information in the nucleosome environment.

Results

Defined genome-wide chromatin reconstitution system with varying nucleosome densities. To assess array generation by remodelers in a biochemically defined way, we used our genome-wide chromatin reconstitution system with purified components (Figure 1A, (Krietenstein et al., 2016)) including recombinant INO80 complex (accompanying paper Oberbeckmann & Krietenstein et al.) and recombinant Chd1 (Farnung et al., 2017). Briefly, genomic plasmid libraries were reconstituted with *Drosophila* embryo histone octamers into nucleosomes by salt gradient dialysis (SGD). SGD chromatin was incubated with ATP,

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

purified yeast remodelers (Figure S1A), and the barrier Reb1 or the restriction enzyme BamHI, which generates double strand breaks (DSBs) that also amount to nucleosome positioning barriers (accompanying paper Oberbeckmann & Krietenstein et al.). Resulting nucleosome patterns were analyzed by MNase-seq. The effective histone-to-DNA mass ratio during SGD was varied from 0.2 to 0.8 yielding low, medium and high nucleosome densities reflected in increasingly extensive MNase-ladders at the same MNase digestion conditions (Figure 1B). Nucleosome density variation was instrumental to distinguish if linker lengths and phasing distances depended on nucleosome density and/or remodeler features.

INO80, ISW2, ISW1a and Chd1, but not Fun30 align regular arrays at the barrier Reb1. We tested all yeast remodelers with known spacing activity, INO80, ISW2, ISW1a and Chd1 (Krietenstein et al., 2016; Lusser et al., 2005; Stockdale et al., 2006; Torigoe et al., 2013; Tsukiyama et al., 1999; Udugama et al., 2011) as well as the Fun30 remodeler, for which it was unclear if it has spacing activity (Awad et al., 2010). INO80, ISW2, ISW1a and Chd1, each in combination with Reb1, generated phased regular arrays at promoters with Reb1 sites (red shaded top of heat maps in Figure 1C), while Fun30 did not (Figure S1B). This clarifies that Fun30 does not have regular array generation and alignment activity.

Previously, Chd1 purified from budding yeast did not show much effect in genome-wide reconstitutions (Krietenstein et al., 2016). This was maybe due to full-length Chd1 tending to aggregate *in vitro*, which is why truncated Chd1 constructs were often used (McKnight et al., 2011; Patel et al., 2011). Here, we leveraged our finding that recombinant full-length Chd1 is stabilized in complex with recombinant FACT complex (Farnung et al., 2017) and achieved *in vitro* array generation and alignment also by Chd1.

The heat map patterns (Figure 1C) and even more the corresponding composite plots for the Reb1-bound genes only (Figure 1D) suggested that the distance of arrays to the barrier Reb1 as well as the linker lengths varied with nucleosome density in a remodeler-specific way. For all remodelers with spacing activity, array extent increased with growing density, consistent with greater nucleosome availability and processive spacing activity. Array extent at high density was larger than in our previous reconstitutions (Krietenstein et al., 2016), i.e., we achieved higher densities here. Adding more remodeler after half of the incubation time did not change the array distances of resulting patterns confirming non-limiting remodeling activity and steady state conditions (Figure S1C).

Remodelers set phasing and spacing distances symmetrically around barriers. To better assess distances to barrier (phasing) and linker lengths (spacing), we aligned the MNase-seq data for each remodeler/barrier/density combination to either *in vivo* Reb1 sites or BamHI sites (Figure 2A). For each replicate (Figure S2A-C), we called nucleosome peaks and determined the distances to barrier and linker lengths as defined in Figure 2B.

All remodelers symmetrically aligned regular arrays to BamHI sites, which are palindromic and therefore inherently symmetrical, and most of them also to Reb1 sites (Figures 2A, S2A,B) regardless of site orientation and position relative to genes (groups 1 to 3; Figure S3A,B). However, if INO80 aligned arrays at promoter Reb1 sites (groups 1 to 3, Figure S3A, accompanying paper by Oberbeckmann & Krietenstein et al.), nucleosome occupancy (peak height) was higher over genic versus non-genic regions at low and medium nucleosome density leading to asymmetric patterns with regard to peak heights in groups 1 and 2. Reb1 site orientation had no effect (group 1 vs. 2). This asymmetry in nucleosome occupancies reflected that positioning of +1 nucleosomes, per definition the first nucleosomes downstream of transcription start sites, i.e. at gene starts, was not only guided by Reb1 bound to promoter sites but also synergistically by underlying DNA shape features (accompanying paper Oberbeckmann & Krietenstein et al.). We recapitulated here that INO80 was able to position *in vivo*-like +1 nucleosomes in the absence of a barrier at low and medium densities (Figure S2C,D). This synergism between Reb1- and DNA shape-guided +1 positioning at low and medium density resulted in higher occupancy at the +1 nucleosomes, which are alignment points for +2 nucleosomes and so on. Therefore, all array peaks over genes were higher than their counterparts over non-genic regions.

However, such synergism was not seen at high density where *in vivo*-like +1 nucleosomes positioning by INO80 alone was much less pronounced (Figure S3C,D). This inability was not due to a general inability of INO80 to slide densely packed nucleosomes as INO80 could generate Reb1-aligned arrays at these high nucleosome densities, too (Figures 1C,D, 2A, S2A,B). Nonetheless, this activity was apparently incompatible with or dominant over DNA shape-guided nucleosome positioning (see Discussion). This showed again that our here generated high nucleosome density was higher than the nucleosome density used previously (Krietenstein et al., 2016), otherwise *in vivo*-like +1 nucleosome positioning by INO80 would not have been clearly observed in our earlier study.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

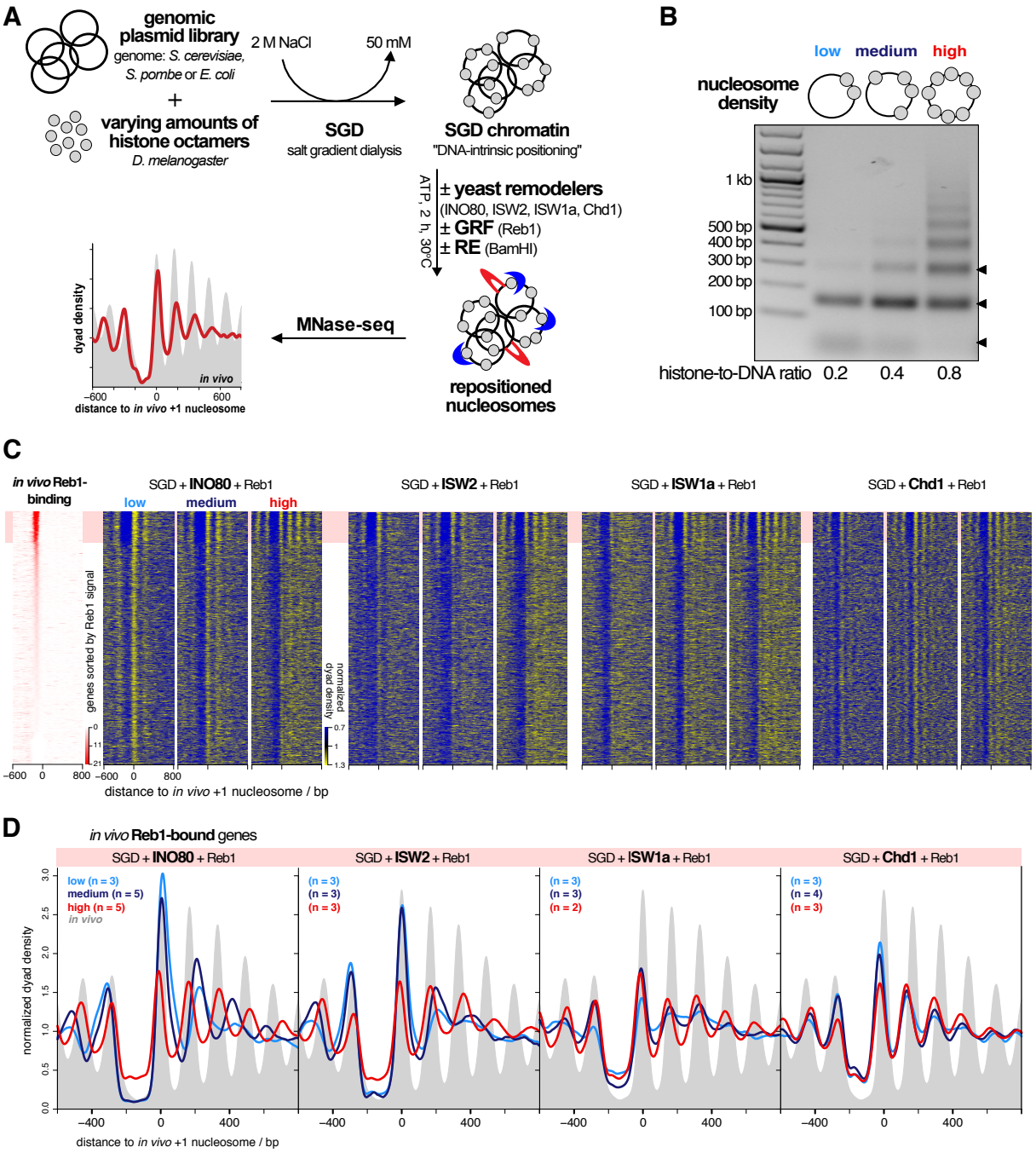


Figure 1. Reb1-guided nucleosome positioning *in vitro* by individual remodelers at varying nucleosome density. (A) Overview of genome-wide *in vitro* reconstitution system. **(B)** Comparison of SGD chromatin reconstituted at indicated histone-to-DNA ratios. DNA fragments after MNase-digest under the same conditions were resolved by agarose gel electrophoresis. Arrow heads on right point to subnucleosomal, mono- and dinucleosomal fragments (bottom to top). **(C)** Heat maps of MNase-seq data for SGD chromatin of the indicated nucleosome density and after incubation with indicated remodelers and Reb1. Chd1 refers to the Chd1/FACT complex. Heat maps are aligned at *in vivo* +1 nucleosome positions and sorted according to decreasing (top to bottom) anti-Reb1 SLIM-ChIP score (Gutin et al., 2018) shown in leftmost heat map. Horizontal red shading highlights genes with strong *in vivo* Reb1-binding in their promoters. Merged data of replicates are shown, individual replicates in Figure S1B,C. **(D)** Composite plots for MNase-seq data averaged over the indicated number of replicates (n) as in panel B but only for genes highlighted in red in panel B.

In this context, we also tested if Fun30 positions *in vivo*-like +1/-1 nucleosomes on its own, but it did not (Figure S3D).

In contrast to nucleosome peak *heights*, nucleosome peak *positions* and therefore corresponding phasing and spacing distances were not significantly affected across groups 1 to 3 for all remodelers, including INO80 (Figure S3B). Therefore, all remodelers symmetrically generated phasing and spacing distances at Reb1 and BamHI sites, which warranted averaging over the up- and downstream values. Resulting values were plotted in different ways to

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

facilitate multi-dimensional comparisons (Figure 2C-E). As all remodelers generated linker lengths independently of the barrier type, we combined linker length values for both barriers (Reb1 and BamHI, Figure 2C). Linker length determination relied on nucleosome peak calling, which was often not possible beyond the -1/+1 nucleosomes at low nucleosome density (Figure S2A), so that linker length data for low density conditions were more sparse, even absent for ISW1a.

Remodeler-specific rulers set spacing in a density-independent or –dependent way. To compare spacing generated by different remodelers at different nucleosome densities, we focused on the averaged length of linker 1 (Figure 2B), which was most accessible across all nucleosome densities. Chd1 generated the shortest (12-13 bp) and ISW1a a bit longer (21-26 bp) linker 1 lengths without significant effects by nucleosome density (Figure 2D,E). ISW2 generated rather constant spacing (54-58 bp) at low and medium but tighter spacing (38 bp) at high density. For INO80, linker lengths steadily increased with decreasing density from 33 to 82 bp.

We concluded that linker lengths and their dependencies on nucleosome density were remodeler-specific and interpreted this as follows. Spacing activity of a remodeler has two aspects. On the one hand, the remodeler equalizes linker lengths leading to regularity in arrays, which is the classical definition of spacing activity (Ito et al., 1997; Varga-Weisz et al., 1997). On the other hand, the resulting linkers have a certain length. In our purified system, this may either be determined by nucleosome density and/or by a remodeler-intrinsic feature. Following (Yamada et al., 2011), we call a remodeler feature that sets nucleosome spacing a "ruler". We use this term also for the feature that sets the distance to barriers (see below). Indicative for a remodeler ruler is remodeler-specific clamping, i.e., if constant spacing is generated at different nucleosome densities (= clamping) and different remodelers generate different spacing (= remodeler-specific), which shows that spacing depends on remodeler-intrinsic and not nucleosome-intrinsic properties (Lieleg et al., 2015). We saw remodeler-specific clamping for Chd1 at all, for ISW1a at high versus medium and for ISW2 at medium versus low densities (Figure 2C-E). As none of the remodelers with spacing activity can disassemble nucleosomes (Clapier and Cairns, 2009) and thereby affect nucleosome density, their rulers can only set their respective linker lengths if these are shorter than or equal to the density-determined linker length at equidistant nucleosome distribution. Accordingly, Chd1 and ISW1a set their ruler-specified linker lengths at all and ISW2 at medium and low densities. ISW2 had to generate shorter linkers at high density and INO80 either did not have a ruler or the ruler responded to changes in nucleosome density.

In vitro mononucleosome assays suggested that INO80 requires at least 40 bp of nucleosome-free DNA for nucleosome sliding (Zhou et al., 2018), while it generated 30 bp linkers in tri-nucleosomes (Udugama et al., 2011). Here, at high nucleosome density, INO80 generated linkers of about 33 bp consistent with previous observations. We tried to enforce even tighter spacing by increasing nucleosome density. This did not decrease spacing and phasing distances but peak heights (Figure S2B,C), probably due to increased aggregation without effective increase in nucleosome density of soluble chromatin.

Remodeler type, barrier type and nucleosome density determine distance to barrier. The findings for the distance to barrier were more complex than for lengths of linker 1 (Figure 2C-E). First, the distance to barrier depended on the barrier type (Figure 2C). It was always longer for Reb1 than for BamHI generated DNA ends, with the largest difference for ISW1a and the smallest for Chd1. The DNA footprint size of *S. cerevisiae* Reb1 is not known, possibly 20 bp as for the *S. pombe* Reb1 DNA binding domain (Jaiswal et al., 2016). This would contribute 10 bp to the distance to barrier (Figure 2B) and could explain the differences between distance to Reb1 vs. BamHI sites for Chd1, but not for the other remodelers. Therefore, INO80, ISW2 and ISW1a, but not Chd1, aligned nucleosomes differently at Reb1 versus at DSBs.

Second, the distance to DNA ends was mostly similar to linker lengths for INO80, ISW2 and ISW1a, arguing that these remodelers, but not Chd1, used a DNA end in a similar way as a neighboring nucleosome for nucleosome alignment.

Third, distances to barriers depended on nucleosome density in a similar way as linker lengths for all remodelers but INO80, where distances to both barriers varied less between low and medium density than linker length.

We concluded that there are remodeler-specific differences in how a nucleosome is positioned next to another nucleosome versus next to a barrier like Reb1 versus next to a DNA end and how this depends on nucleosome density. This is again a clear case of different remodelers generating different nucleosome positioning, although starting from the same SGD chromatin, which argues for remodeler-specific rulers governing nucleosome positioning.

Remodelers differ in processivity of nucleosome positioning. All remodelers generated similar lengths of linker 1 to linker 3 at high density (Figure

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

2D), which we interpreted as processive spacing activity along the arrays as long as nucleosomes were sufficiently provided. At low density, ISW2, Chd1 and especially INO80 still generated high +1/-1 nucleosome peaks (Figure S2A), in contrast to ISW1a, for which these peaks were less pronounced and +2/-2 nucleosome peaks could not be discerned. We suggest that ISW1a is less processive than other remodelers in bringing nucleosomes next to barriers at low densities.

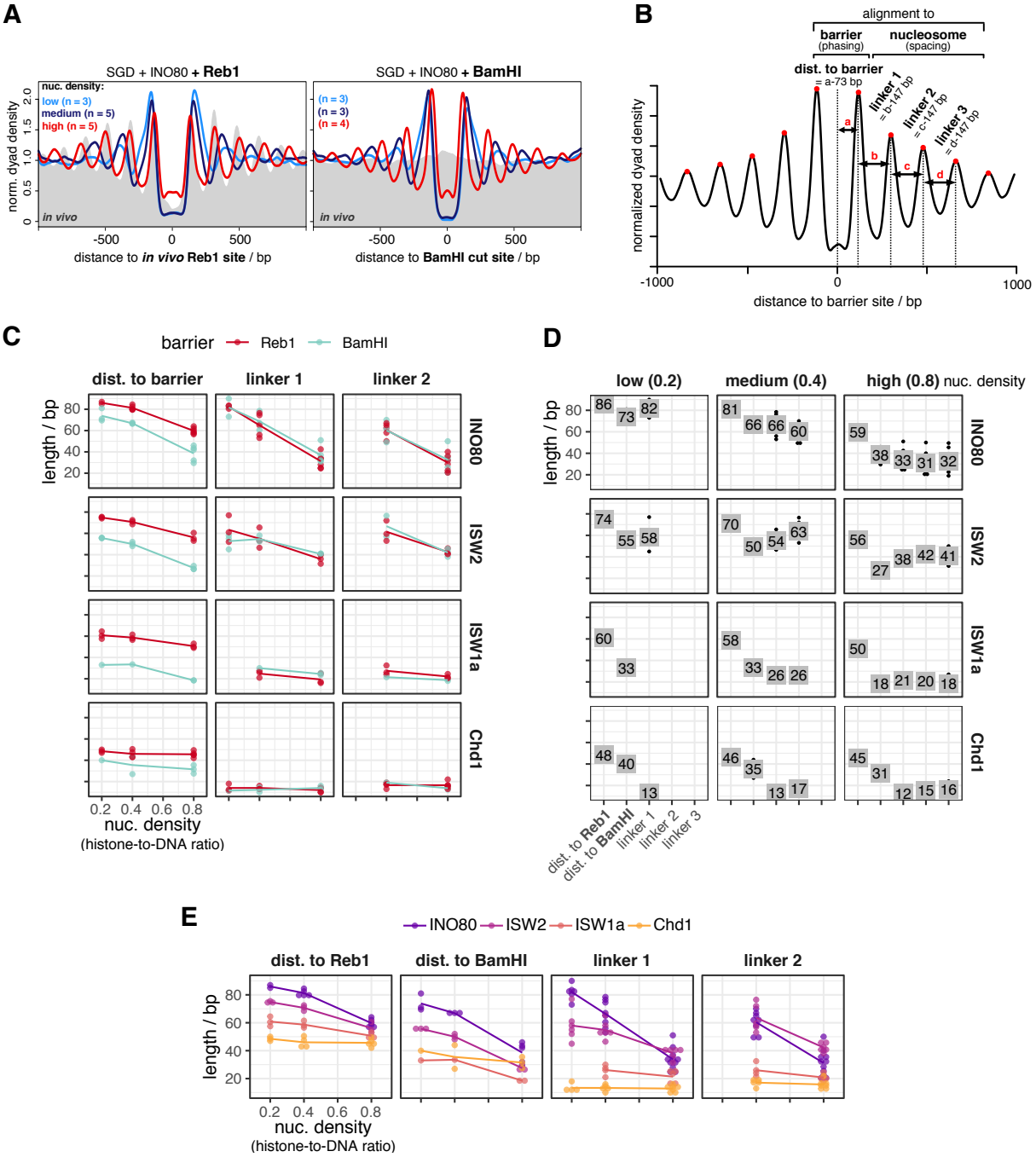


Figure 2. Quantification of barrier-aligned nucleosome array features depending on barrier, remodeler and nucleosome density. (A) Composite plots of same MNase-seq data for INO80 as in Figure 1D but aligned at anti-Reb1 SLIM-ChIP-defined Reb1 sites (left), or at BamHI sites (right) of SGD chromatin reconstituted at the indicated nucleosome densities and incubated with INO80 and BamHI. **(B)** Scheme defining array features quantified from barrier-aligned composite plots as in panel A. **(C) – (D)** Array feature values for the indicated combinations of barrier, remodeler and nucleosome density plotted in different ways allowing comparison between barriers (especially panel C), values (especially panel D) and remodelers (especially panel E). Chd1 refers to the Chd1/FACT complex. Panel D and Figure S2A-C show individual replicates, panels C and E replicate averages.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

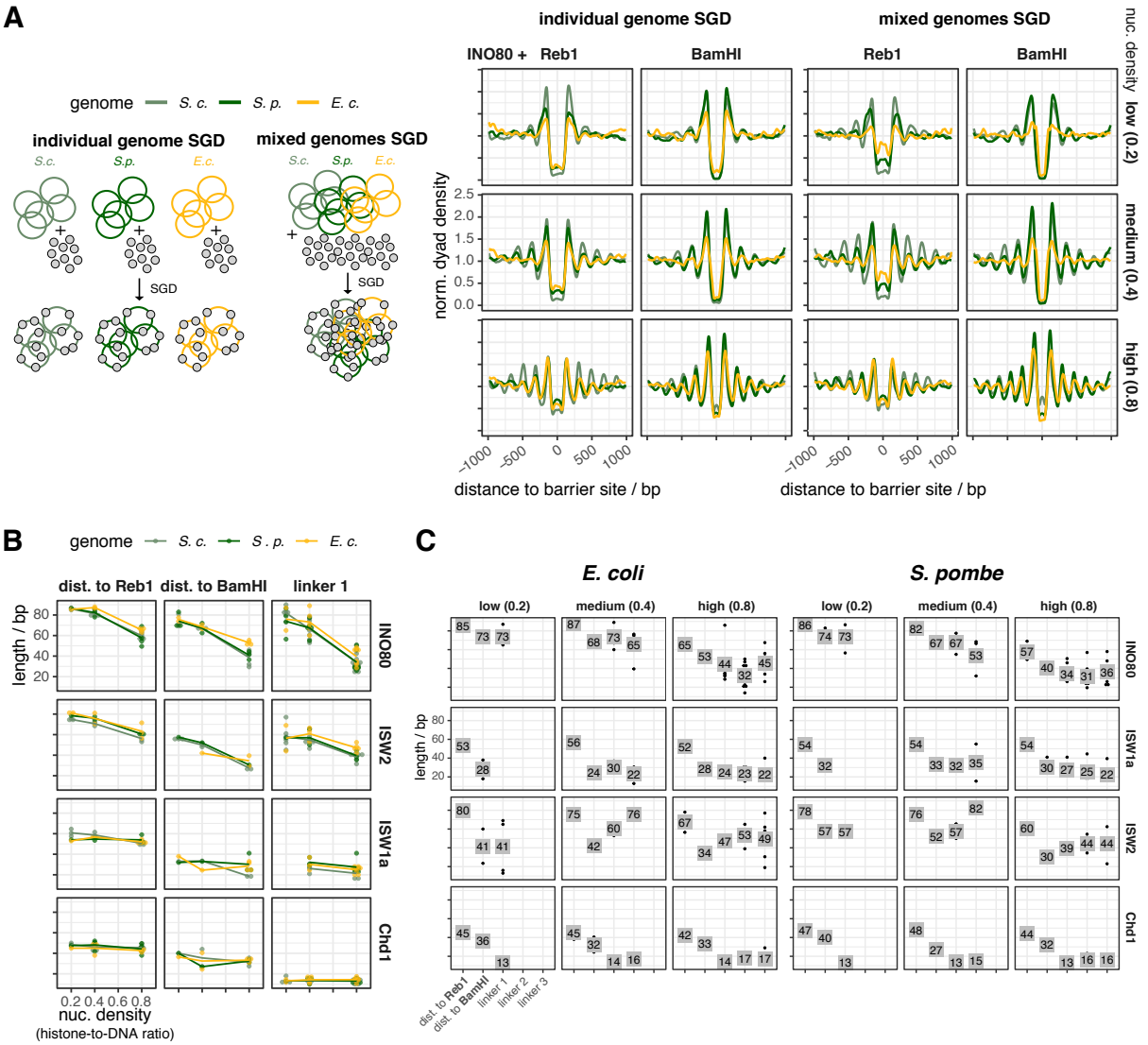


Figure 3. Yeast remodelers generate arrays on heterologous genomes with same spacing and phasing distances as on the yeast genome. (A) Left: Schematic showing SGD reconstitution with individual or mixed genomes. Right: Composite plots as Figure 2A but for the indicated barriers and either individual (left) or mixed (right) genomes. Reb1 sites were called by PWM. **(B)** and **(C)** As Figures 2C,D, respectively, but for the indicated genomes. Individual replicates in Figure S4.

Remodelers generate similar arrays on all but more effectively on eukaryotic DNA sequences. The same linker lengths in arrays at BamHI and Reb1 sites (Figure 2C), at Reb1 sites in groups 1 to 3 and the symmetry of nucleosome distances to Reb1 sites in groups 1 to 3 (Figure S3A,B) suggested that remodeler rulers position nucleosomes independently of DNA sequence flanking the barriers. Nonetheless, there are evolved DNA features at promoters, especially for INO80 (accompanying paper Oberbeckmann & Krietenstein et al.), that affected occupancies (peak heights, not positions, Figure S3A), which may also be true for evolved nucleosome-favoring dinucleotide periodicities (Satchwell et al., 1986) in gene bodies.

To rigorously disentangle these contributions, we tested the remodeler/barrier/density combinations also with SGD chromatin of *S. pombe* and *E. coli* genomic plasmid libraries (Figures 1A, 3A,B, S4A), including the steady state control (Figure S4B). We did not observe substantial differences in spacing/phasing distances on these genomes for all remodelers, but some replicates, especially at medium and low density, showed lower relative occupancies for the *E. coli* genome.

We concluded that all remodelers align arrays at Reb1 or DSBs regardless of the underlying sequence. Nonetheless, they are more effective in terms of relative occupancies on eukaryotic genomes, likely due to dinucleotide periodicities (Zhang et al., 2009).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

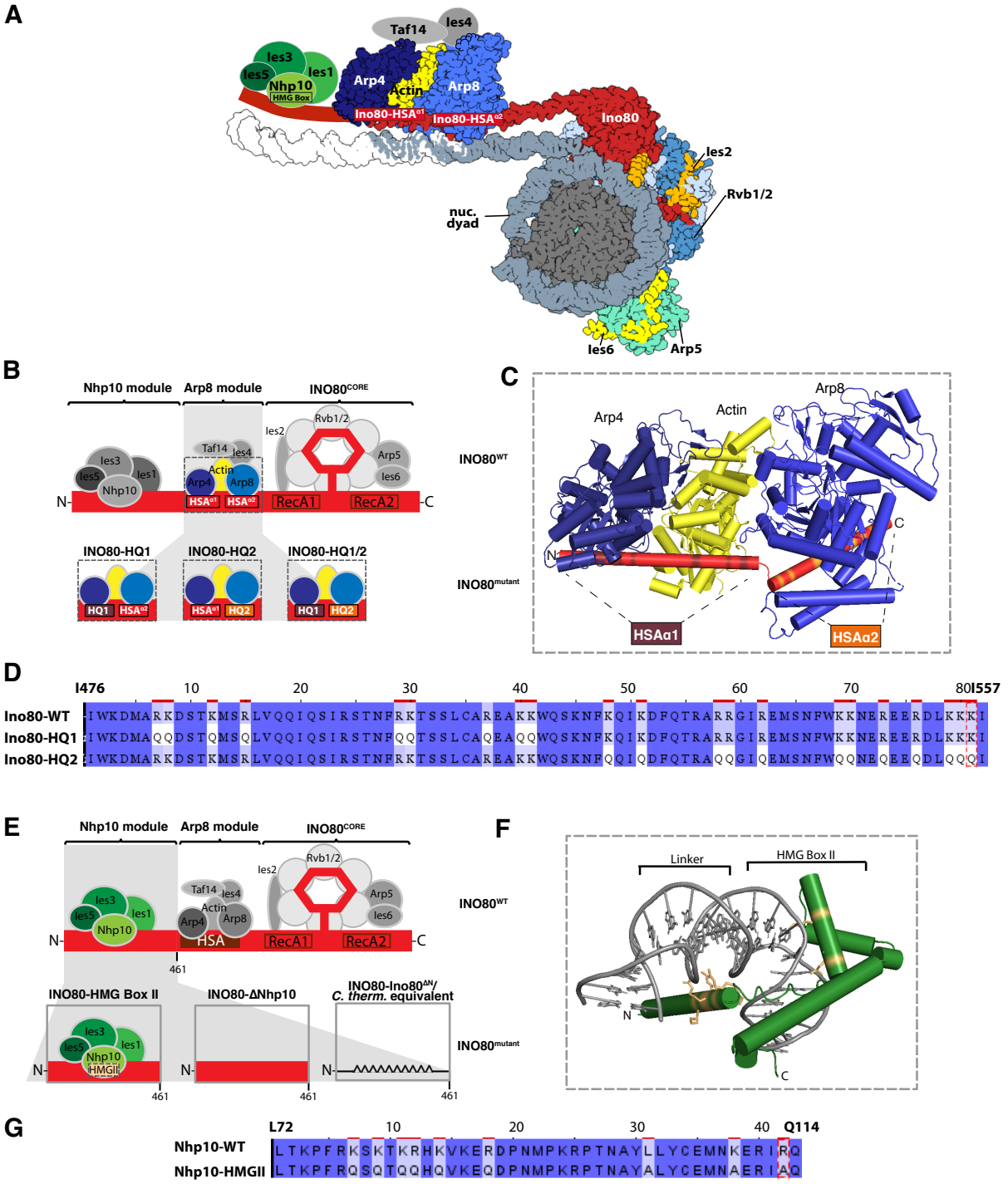


Figure 4. Construction of INO80 mutant complexes. (A) Structure-based (Eustermann et al., 2018; Knoll et al., 2018) model of a nucleosome bound by the INO80 complex with indicated subunits. Nhp10 module, Taf14 and Ies4 organization is assumed. **(B)** Schematic of INO80 complex submodule and subunit organization (top). Zoom into Arp8 module showing three mutant versions (bottom). **(C)** Cylindrical representation of the Arp8 module structure showing mutated residues of Ino80 HSA domain (highlighted in brown and orange). **(D)** Sequence alignment showing mutated residues in Ino80-HQ1 and –HQ2 mutants. **(E)** Schematic of INO80 complex organization as in panel B (top) but zoom into Nhp10 module (bottom) showing three mutant versions. **(F)** Model of Nhp10 HMG box-like and Linker region (residues 62-172) based on TFAM structure (pdb 3tq6). **(G)** Sequence alignment showing mutated residues in Nhp10-HMGII mutant.

INO80 complexes mutated in the Arp8 and/or Nhp10 module. It was unexpected that the clamping criterion did not clearly show a ruler for INO80 (Figure 2C-E), because the INO80 structure suggested modules that bind extranucleosomal DNA and could serve as ruler (Knoll et al., 2018). To clarify, we took advantage of the biochemical accessibility of our recombinant INO80 preparation, the modular

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

INO80 composition and the high-resolution structures (Eustermann et al., 2018; Knoll et al., 2018) to generate candidate mutations that may tune and thereby reveal INO80's ruler.

The INO80 complex has two modules with a likely role in ruler function. First, the Arp8 module consisting of N-Actin, Arp8, Arp4, Taf14 and Ies4 (Figure 4A). It binds the Ino80 main ATPase HSA domain, which is structured as a long helix with a kink that subdivides it into the HSAα1 and HSAα2 part (Knoll et al., 2018). Both bind to extranucleosomal DNA, and mutating DNA contacting lysine residues in HSAα1 or HSAα2 to glutamines (HQ1 and HQ2 mutant, respectively, Figure 4B,C,D) impaired, and combining both mutations (HQ1/2 mutant) abolished mononucleosome centering activity (Knoll et al., 2018).

The second, Nhp10 module, binds the Ino80 ATPase N-terminus, and contains the HMG box Nhp10 subunit, along with Ies1, Ies3 and Ies5 (Figure 4A,E). This module is species-specific and affects the processivity and extranucleosomal DNA requirements in mononucleosome sliding assays (Zhou et al., 2018).

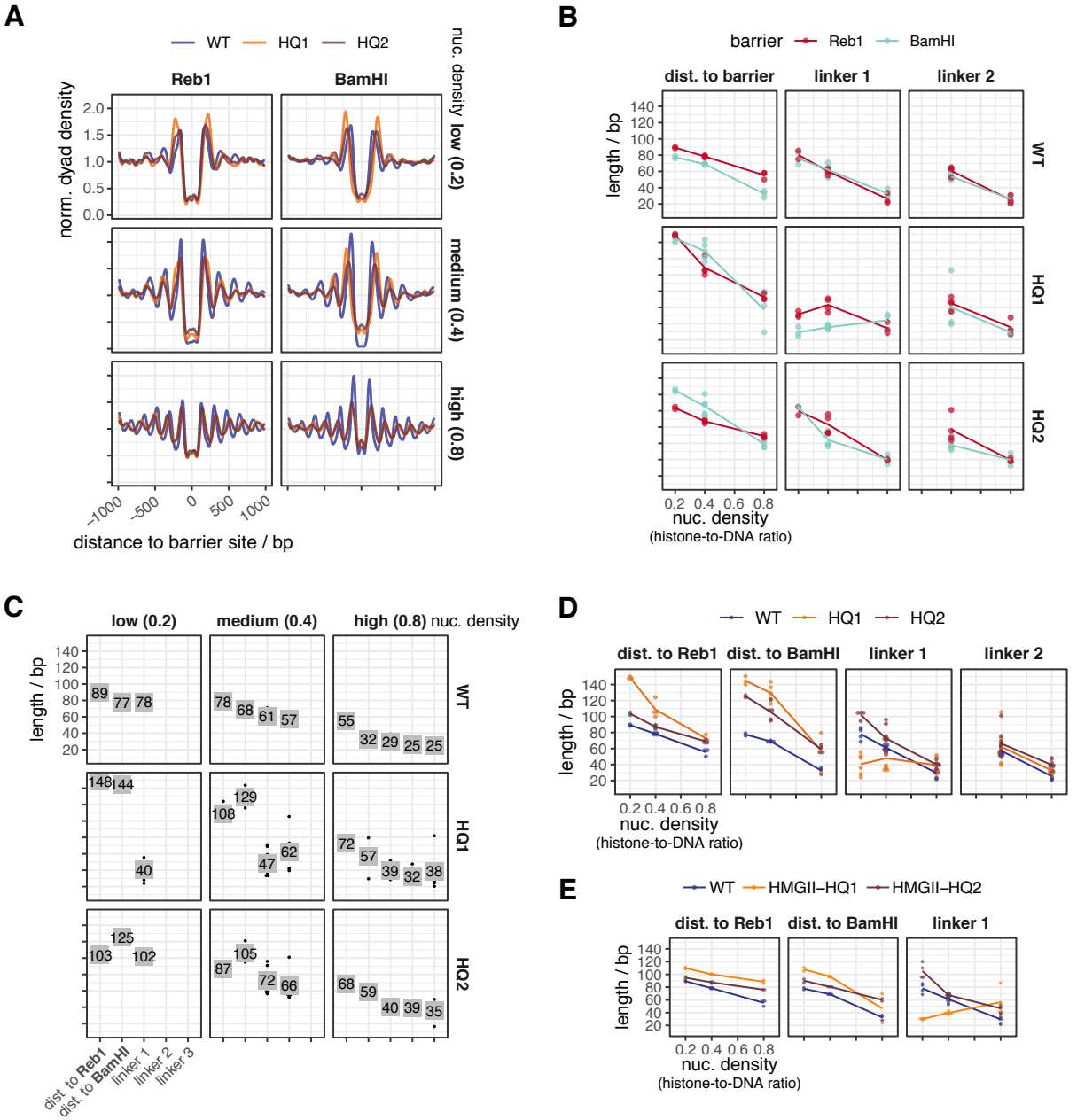


Figure 5. Mutations in the INO80 Arp8 module affect the generation of array features. (A) Composite plots as in Figure 2A but for the indicated WT and mutant INO80 complexes and nucleosome densities**. (B) – (D)** As Figure 2C-E, but comparing indicated WT and mutant INO80 complexes. Individual replicates in Figures S5A,B. **(E)** As Figure 2E, but for the indicated WT and mutant INO80 complexes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

Calculating a homology model for Nhp10 based on another HMG box protein, TFAM (Ngo et al., 2014), we inferred and mutated amino acid residues putatively involved in Nhp10-DNA interactions (HMGII mutant, Figure 4F,G). These mutations were also combined with the HQ1 or HQ2 mutants (HMGII-HQ1 and HMGII-HQ2). Further, we prepared recombinant INO80 complex without any Nhp10 module subunits (ΔNhp10 mutant, no truncation of the Ino80 ATPase N-terminus) or a version where the Ino80 ATPase lacked residues 1-461 (INO80$^{\Delta N}$ mutant), which removes the assembly platform for the Nhp10 module (Figure 4A).

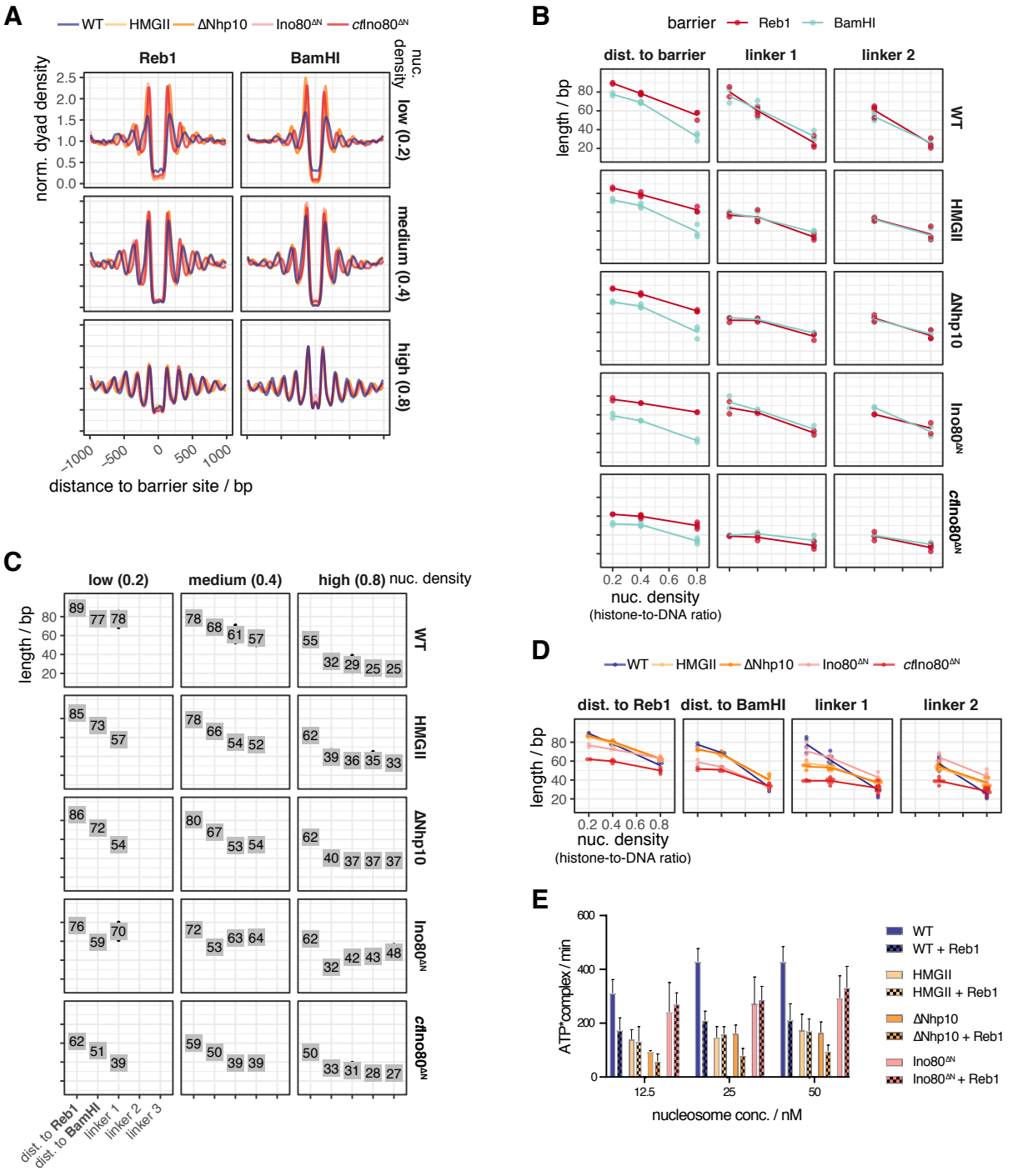


Figure 6. Mutations in the INO80 Nhp10 module affect the generation of array features. (A) – (D) As Figure 5A-D but for the indicated WT and mutant INO80 complexes. Individual replicates in Figure S5A. **(E)** Mononucleosome-stimulated ATPase activities for the indicated nucleosome concentrations and INO80 complexes, and respective equimolar Reb1 concentrations. The extranucleosomal DNA of the 601 mononucleosome contained a Reb1 site at 70 bp distance from the 601 sequence and Reb1 was included as indicated.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

INO80 mutant complexes reveal a multilayered ruler. All mutant complexes were assayed like the wild type (WT) INO80 complex (Figures 5A-E, 6A-D, S1A, S5A,B). WT INO80 was assayed again alongside with matching SGD chromatin. Comparing these replicates (Figure 5C) with previous values for WT INO80 (Figure 2D) reflected variability in preparing SGD chromatin but at the same time the robustness of the overall effects. All tested INO80 mutants generated steady-state patterns (Figure S5B) and differed from WT INO80 in forming aligned arrays in the following ways.

First, all mutants, besides the HQ1/2 mutant, which was almost inactive (Figure S5A), as expected (Knoll et al., 2018), generated phased regular arrays, but with varying effectiveness and altered distance to one or both barrier types and/or linker lengths compared to WT INO80 (Figures 5D,E, 6D). This revealed that also INO80 has a ruler, to which both the Arp8 and the Nhp10 module contribute.

Second, the HQ1 showed stronger effects than the HQ2 mutation (Figure 5D). Both increased the distances to both barriers. While HQ2 increased linker length at all densities, HQ1 gained clamping activity, i.e., linker length hardly depended on nuclesome density. Both mutations uncoupled distance to DNA ends from linker lengths, in contrast to WT INO80 (Figure 2D,E). Only for HQ1, linker 1 length depended on barrier type (Figure 5B). We concluded that the Arp8 module, especially via HSAα1 helix-DNA interactions, is threefold involved in spacing, alignment to barrier and responding to nucleosome density.

Third, the Nhp10 module subunits contributed to the ruler mainly through the HMG box of Nhp10 as the respective point mutations (HMGII mutant) mimicked the effects upon lack of all Nhp10 module subunits (ΔNhp10 mutant) (Figure 6C,D). With these mutations, distances to both barriers were not much affected, but linker length depended less on density, i.e., clamping was gained, similar to the HQ1 mutation. Effects of the combined HMGII-HQ1 and –HQ2 mutations were dominated by the HQ mutations, but with reduced effects on distance to barriers (Figure 5E). Even though the Nhp10 HMG box was a prime candidate for sensing extranucleosomal DNA, its contribution was minor compared to the HSA helix contribution.

Fourth, the INO80$^{\Delta N}$ mutation affected the distance to Reb1 and even more to DNA ends, but gained clamping less strongly than the HMGII or ΔNhp10 mutations (Figure 6D). The INO80$^{\Delta N}$ mutant lacked the complete Nhp10 module, but also the Ino80 ATPase N-terminus and Taf14 (Figure S1A), which may account for the differential effects.

Fifth, the INO80$^{\Delta N}$, HQ1 and HQ2 mutations most drastically affected distance to BamHI sites, but in opposite ways (Figures 5C,D, 6C,D).

Effects on nucleosome stimulated ATPase activity versus on ruler function are not strictly coupled. The WT INO80 ATPase activity is stimulated by nucleosomes and inhibited about twofold in the presence of Reb1 (accompanying paper Oberbeckmann & Krietenstein et al.). This relative inhibition by Reb1 was not seen or less pronounced for the mutated INO80 complexes (Figure 6E). The ATPase activity of HMGII and ΔNhp10 mutants was similar to that of WT INO80 in the presence of Reb1. The INO80$^{\Delta N}$ mutant had intermediate activity. We concluded that all tested mutants were affected both with regard to ATPase activity and with regard to their ruler but that both effects were not strongly coupled.

***Chaetomium thermophilum* INO80 core complex suggests species-specific ruler.** The INO80 core complex of *C. thermophilum*, which we previously used for cryoEM studies (Eustermann et al., 2018), corresponds to the *S. cerevisiae* INO80$^{\Delta N}$ mutant as it also lacks its Ino80 ATPase N-terminus. It showed stronger clamping and generated shorter linkers and distances to Reb1 than INO80$^{\Delta N}$ at all densities, and much shorter linkers and distances to both barriers than *S. cerevisiae* WT INO80 at low and medium densities (Figure 6B-D). This suggests that INO80's ruler may be species-specific.

Discussion

Our study answers one of the oldest questions in chromatin research: what determines the spacing and phasing distances of nucleosome arrays in absolute terms? The solution to this question are ATP dependent remodelers from the ISWI, CHD and INO80 families with spacing activity. These do not only equalize linker lengths but, as we reveal here, bear rulers for setting distances between two adjacent nucleosomes and between nucleosomes and other alignment points.

Remodeler rulers explain previous *in vivo* observations. Rulers combined with barriers mechanistically explain *in vivo* observations that involved ISW1a, ISW2, Chd1 and INO80 in +1 nucleosome positioning and/or array regularity and phasing (Gkikopoulos et al., 2011; Hennig et al., 2012; Kubik et al., 2019; Ocampo et al., 2016; Parnell et al., 2015; Pointner et al., 2012; van Bakel et al., 2013; Whitehouse et al., 2007; Yen et al., 2012).

The average *S. cerevisiae* linker length of 18 bp (Thomas and Furber, 1976) results from combined contributions of ISW1a and Chd1 (Ocampo et al., 2016). As we show that ISW1a and Chd1 rulers

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

generate linkers of about 20 and 12 bp, respectively, the 18 bp average linker speaks for ISW1a contributing globally more than Chd1. Indeed, lack of Isw1 *in vivo* globally shortened linkers, while lack of Chd1 affected global spacing only mildly (Kubik et al., 2019; Ocampo et al., 2016). Locally, high transcription rate correlates with shorter spacing (Chereji et al., 2018; Ocampo et al., 2016), which points to increased Chd1 contribution, probably due to increased Chd1 recruitment by elongating RNA polymerase (Simic et al., 2003).

Remodeler-specific rulers explain how ISW1a, ISW2 and INO80 affect +1 nucleosome positioning *in vivo* (Kubik et al., 2019; Parnell et al., 2015; Whitehouse et al., 2007; Yen et al., 2012) and *in vitro* (Krietenstein et al., 2016), especially in combination with RSC. RSC and SWI/SNF are the only yeast remodelers that disassemble nucleosomes (Clapier and Cairns, 2009; Clapier et al., 2017), particularly at promoter NDRs (Badis et al., 2008; Brahma and Henikoff, 2019; Ganguli et al., 2014; Hartley and Madhani, 2009; Kubik et al., 2019; Kubik et al., 2018; Parnell et al., 2008; Rawal et al., 2018; van Bakel et al., 2013; Wippo et al., 2011). By definition, a promoter NDR has low nucleosome density. Therefore, remodeler rulers will set distances to NDR-bound barriers as measured here at low or medium nucleosome density. *In vivo* distances between Reb1 and +1 nucleosomes are 60-80 bp (Figure S3B, (Rhee and Pugh, 2011)), which are within remodeler-specific distances to Reb1 at medium or low density (81-86 bp for INO80, 70-74 bp for ISW2, 58-60 bp for ISW1a). ISW2 and INO80 contribute more to +1 nucleosome positioning *in vivo* than ISW1a (Kubik et al., 2019) as their long rulers are more suited for setting long distances across NDRs. Conversely, the short Chd1-ruler hardly contributes to +1 positioning *in vivo* (Kubik et al., 2019; Ocampo et al., 2016; van Bakel et al., 2013). These different ruler characteristics explain why ISW1a and Chd1 are mainly involved in spacing nucleosomes into densely packed arrays and why ISW2 and INO80 mainly use their ruler for +1 alignment at NDRs *in vivo*. This resolves the conundrum (Krietenstein et al., 2016) why yeast has two remodelers, INO80 and ISW2, that seemingly generate "too wide" spacing compared to average *in vivo* spacing. We do not preclude that other mechanisms, like recruitment via histone modifications or transcription factors, also affect where each remodeler is active.

Functional and structural identification of remodeler rulers. The protein ruler model was first proposed for ISW1a (Yamada et al., 2011). It suggested that ISW1a shortens the linker until its ruler contacts the neighboring nucleosome, but did not conceptualize why this would lead to a stable nucleosome position. We built on and expanded this model, identified remodeler rulers via their functionality and pinpointed the INO80 ruler also in structural terms. On the functional level, a ruler is revealed if

a) the same remodeler generates the same phasing and/or spacing distances although it works on chromatin with varying nucleosome density (clamping activity), or
b) different remodelers/different mutant versions of the same remodeler generate different phasing and/or spacing distances although they all work on the same chromatin (remodeler-specific phasing/spacing).

For the INO80 complex, we found that the Nhp10 module, especially the Ino80 N-terminus, as well as the Arp8 module, especially the Ino80-HSA-helix, contributed to the ruler function. Lack of the Ino80 N-terminus, concomitant with lacking the Nhp10 module, allowed INO80, e.g., to slide nucleosomes closer to DNA ends, maybe for steric reasons, while impaired DNA traction during remodeling due to compromised Ino80-HSA helix-DNA interactions had the opposite effect. It remains to be elucidated how exactly such modules within the multi-subunit organisation relay barrier information to the core ATPase.

Remodeler rulers regulate nucleosome sliding direction bias in response to nucleosome environment. We propose an overarching framework for this relay that amounts to a widely applicable remodeler ruler principle (Figure 7). A remodeler may slide a nucleosome either to the left or to the right from a given position. If there is no bias for sliding in either direction, the nucleosome will experience a random walk along the DNA (regions C in in three hypothetical examples Figure 7A). Net nucleosome movement in one direction (Gangaraju and Bartholomew, 2007; Langst et al., 1999; McKnight et al., 2011; Stockdale et al., 2006; Udugama et al., 2011; Yang et al., 2006; Zhou et al., 2018) requires an overall sliding direction bias in this direction. We conceptualize a remodeler ruler as a remodeler-intrinsic feature that generates an overall sliding direction bias in response to the (epi)genetic information in the environment of the nucleosome that the remodeler is remodeling. The bias may originate from differences, e.g., in binding orientation, ATPase activity, sliding rate or processivity and is regulated by interaction of the ruler with a generalized "barrier". This may be a GRF, a DSB, a neighboring nucleosome, or a DNA sequence element. Histone modifications/ variants may modulate as well. While the microscopic details may differ for different remodelers and information

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

input, the overall regulation of sliding direction bias by the ruler will share three key elements that constitute the ruler mechanism. First, the ruler has a certain reach (regions A + B in Figure 7A), within which it interacts with the barrier. Second, if the position, from where the remodeler slides the nucleosome, is within region B, the interaction between ruler and barrier biases overall sliding direction towards the barrier (red curve is above green curve), e.g., due to binding energy gained upon orienting the remodeler towards vs. away from the barrier. Third, if the nucleosome is in region A, the ruler-barrier interaction disfavors sliding towards relative to sliding away from the barrier (green curve is above red curve), e.g., because the ruler gets sterically in the way. Our study determined the length of region A for different remodeler and barrier types and conditions. Region B and exact curve shapes will have to be determined in future studies. If these three key elements are met, resulting fluxes lead to steady-state nucleosome placement at a defined position relative to the barrier (stippled vertical arrows throughout Figure 7). This position is a self-stabilizing dynamic equilibrium point (intersection of red and green curves) without sliding direction bias here, but with biases *towards* this point from neighboring positions. This model applies to how a remodeler with ruler stably positions a nucleosome next to a GRF as well as to another nucleosome and therefore explains both spacing and phasing.

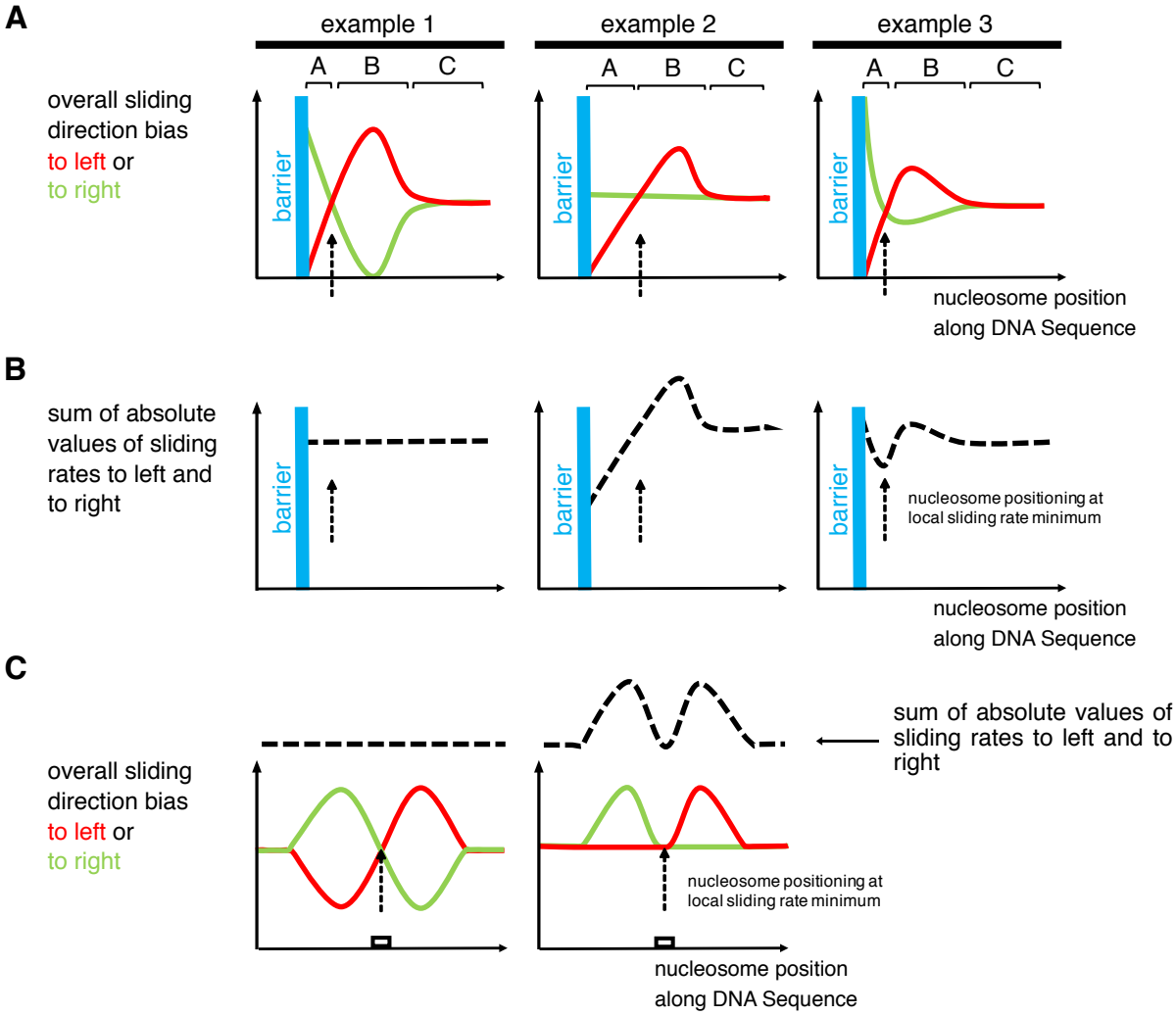


Figure 7. Model for remodeler ruler mechanism. (A) Three hypothetical examples for how a remodeler ruler regulates the overall bias of sliding a nucleosome to left (red curves) or to right (green curves) resulting in nucleosome positioning (stippled vertical arrows) in the vicinity of a barrier. **(B)** As panel A but plotting sum of absolute values of sliding rates to the left and to the right (stippled black curves). **(C)** Two hypothetical examples for how a remodeler ruler leads to nucleosome positioning over a DNA site (white box). Symbolics as in panels A and B. For details see text.

It also explains density-independent clamping. As long as a remodeler is processive enough to fortuitously bring nucleosomes into region B of a barrier also at low density, the ruler mechanism will keep the nucleosome at the dynamic equilibrium point. Nonetheless, the model can also accommodate sensing of nucleosome density and barrier type, e.g., if the ruler offers a hierarchy of interaction points

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

that depends on density or barrier type. For example, INO80 may be able to adopt different conformations that may provide different interaction sites and have different footprint sizes, which may explain why INO80 can remodel arrays with just 30 bp linkers despite a measured footprint of >50 bp (Brahma et al., 2018). INO80 mutants showed not concerted but uncoupled effects on distances to Reb1, DNA ends and nucleosomes, even if the same module, like the Nhp10 module, was differentially mutated. Chd1 generated shorter linker lengths (12-16 bp) than distances to DNA ends or Reb1 (35-40 bp). For Chd1, Reb1 may be a "hard" barrier while nucleosomes are "soft" barriers as they are partially "invaded" by the ruler. Indeed, Chd1 partially unwraps nucleosomal DNA (Farnung et al., 2017). The way how different remodeler rulers interact with different barriers requires clarification, and we outline our model (Figure 7) in terms of extension-less point particles, but actual footprints have to be taken into account. The model is fully compatible with the ruler, i.e., the DNA binding domain (DBD) of Chd1 (McKnight et al., 2011) or *Drosophila* ACF (Yang et al., 2006), introducing bias via sensing extranucleosomal DNA length. Indeed, differently long extranucleosomal DNA in mono- or oligonucleosome sliding assays amounts to different distances to barriers like DNA ends or other nucleosomes. Our model is fully consistent with previous data and models but offers an alternative interpretation and is more widely applicable, e.g., to stable nucleosome positioning at only one barrier and not only in-between two barriers.

We introduced our model in terms of overall sliding direction bias. More specifically, the model may refer to differential regulation of sliding rates, i.e., the y-axis in Figure 7A could correspond to "overall sliding rate to the left or to the right". If sliding rates are reciprocally regulated (example 1, Figure 7A), the sum of absolute sliding rate values is constant at each position (Figure 7B), but not upon asymmetric regulation of sliding direction (examples 2 and 3, Figure 7A). As special case (example 3, Figure 7A,B), the dynamic equilibrium point may correspond to a minimum of absolute sliding rate. This case corresponds to the "kinetic release" model (Manelyte et al., 2014; Rippe et al., 2007), which posits that remodelers position nucleosomes at sites where the nucleosome is the (locally) poorest substrate for remodeling.

Ruler-regulated sliding: the unifying principle for nucleosome positioning by remodelers. As nucleosome positions are defined by the DNA sequence bound by the histone octamer, all mechanisms that generate consistent nucleosome positions across many genome copies, must select certain DNA sequences in competition with other sequences. As shown here and in the accompanying paper (Oberbeckmann & Krietenstein et al.), remodelers may mediate this selection in two ways. On the one hand, a remodeler may directly choose a sequence, e.g., INO80 turns DNA shape features into +1 nucleosome positions at promoters (accompanying paper Oberbeckmann & Krietenstein et al.) On the other hand, a remodeler ruler may place a nucleosome at a ruler-determined distance to a barrier, e.g., ISW2 aligns nucleosomes to Reb1 and generates a regular array by aligning a second nucleosome to the first and so on. In the former case, the resulting nucleosomal sequence is directly selected for its sequence features, while in the latter case, it is indirectly selected without regards for its sequence features but merely for its position relative to the barrier, as we show here by using Reb1 sites in *S. pombe* and *E. coli* genomes.

Our ruler model unifies these positioning mechanisms. The generalized barrier also encompasses DNA sequence elements, with which a remodeler ruler interacts such that sliding direction bias is regulated (Figure 7C). This explains observations for hybrid Chd1 remodelers where the Chd1 DBD was replaced with heterologous sequence-specific DBDs (Donovan et al., 2019; McKnight et al., 2011; McKnight et al., 2016). Such hybrid Chd1 remodelers slide nucleosomes faster towards the cognate site of the heterologous DBD, if it was in reach of this site, until the nucleosome became positioned on the site. In our model, the heterologous DBD is a remodeler ruler. As a DNA sequence element as barrier is no hindrance for nucleosome sliding, the remodeler may slide the nucleosome onto this site. This prevents ruler binding to the site, abolishes the increase in sliding rate linked to ruler binding and makes a nucleosome on the cognate site a poorer nucleosome sliding substrate than at neighboring positions (Figure 7C, right), which corresponds to the kinetic release model as noted (McKnight et al., 2011). Our model now adds that sliding from neighboring positions will always (within ruler reach) convene at the cognate site and stabilize this position, even if there is no local sliding rate minimum, as long as the ruler regulates sliding direction bias according to the three key elements outlined above (Figure 7C, left). As our INO80 mutations differently affected nucleosome positioning via DNA shape (accompanying paper Oberbeckmann & Krietenstein et al.) vs. relative to Reb1 vs. DNA ends vs. nucleosomes, the ruler elements seem to be multilayered and maybe linked to different structural conformations. For example, the INO80 conformation required for aligning nucleosomes at high density may not be compatible

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

with positioning +1 nucleosomes via DNA shape. Nonetheless, as the Ino80-HSA mutations affected nucleosome positioning both via DNA shape (accompanying paper Oberbeckmann & Krietenstein et al.) and relative to barriers, the Ino80-HSA domain is a functionally crucial part of to the INO80 ruler.

In vivo there are many ways that may regulate nucleosome positioning by remodelers, e.g., by recruitment, by architectural factors, by nucleosome density fluctuations or by histone variants and modifications, possibly in the context of elongating polymerases. Nonetheless, we expect that the regulation of nucleosome sliding direction bias via built-in sensing and processing of information in the nucleosome environment, i.e., a remodeler ruler, will be at the heart of each nucleosome positioning mechanism.

References

Awad, S., Ryan, D., Prochasson, P., Owen-Hughes, T., and Hassan, A.H. (2010). The Snf2 homolog Fun30 acts as a homodimeric ATP-dependent chromatin-remodeling enzyme. The Journal of biological chemistry *285*, 9477-9484.

Badis, G., Chan, E.T., van Bakel, H., Pena-Castillo, L., Tillo, D., Tsui, K., Carlson, C.D., Gossett, A.J., Hasinoff, M.J., Warren, C.L., *et al.* (2008). A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. Molecular cell *32*, 878-887.

Baldi, S., Jain, D.S., Harpprecht, L., Zabel, A., Scheibe, M., Butter, F., Straub, T., and Becker, P.B. (2018a). Genome-wide Rules of Nucleosome Phasing in Drosophila. Molecular cell *72*, 661-672.e664.

Baldi, S., Korber, P., and Becker, P.B. (2020). Beads on a string-nucleosome array arrangements and folding of the chromatin fiber. Nature structural & molecular biology *27*, 109-118.

Baldi, S., Krebs, S., Blum, H., and Becker, P.B. (2018b). Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. Nature structural & molecular biology *25*, 894-901.

Brahma, S., and Henikoff, S. (2019). RSC-Associated Subnucleosomes Define MNase-Sensitive Promoters in Yeast. Molecular cell *73*, 238-249.e233.

Brahma, S., Ngubo, M., Paul, S., Udugama, M., and Bartholomew, B. (2018). The Arp8 and Arp4 module acts as a DNA sensor controlling INO80 chromatin remodeling. Nature communications *9*, 3309.

Challal, D., Barucco, M., Kubik, S., Feuerbach, F., Candelli, T., Geoffroy, H., Benaksas, C., Shore, D., and Libri, D. (2018). General Regulatory Factors Control the Fidelity of Transcription by Restricting Non-coding and Ectopic Initiation. Molecular cell *72*, 955-969.e957.

Chereji, R.V., Ramachandran, S., Bryson, T.D., and Henikoff, S. (2018). Precise genome-wide mapping of single nucleosomes and linkers in vivo. Genome biology *19*, 19.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annual review of biochemistry *78*, 273-304.

Clapier, C.R., Iwasa, J., Cairns, B.R., and Peterson, C.L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nature reviews Molecular cell biology *18*, 407-422.

Donovan, D.A., Crandall, J.G., Banks, O.G.B., Jensvold, Z.D., Truong, V., Dinwiddie, D., McKnight, L.E., and McKnight, J.N. (2019). Engineered Chromatin Remodeling Proteins for Precise Nucleosome Positioning. Cell reports *29*, 2520-2535.e2524.

Eustermann, S., Schall, K., Kostrewa, D., Lakomek, K., Strauss, M., Moldt, M., and Hopfner, K.P. (2018). Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. Nature *556*, 386-390.

Farnung, L., Vos, S.M., Wigge, C., and Cramer, P. (2017). Nucleosome-Chd1 structure and implications for chromatin remodelling. Nature.

Flaus, A., Martin, D.M., Barton, G.J., and Owen-Hughes, T. (2006). Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. Nucleic acids research *34*, 2887-2905.

Ganapathi, M., Palumbo, M.J., Ansari, S.A., He, Q., Tsui, K., Nislow, C., and Morse, R.H. (2011). Extensive role of the general regulatory factors, Abf1 and Rap1, in determining genome-wide chromatin structure in budding yeast. Nucleic acids research *39*, 2032-2044.

Gangaraju, V.K., and Bartholomew, B. (2007). Dependency of ISW1a chromatin remodeling on extranucleosomal DNA. Molecular and cellular biology *27*, 3217-3225.

Ganguli, D., Chereji, R.V., Iben, J.R., Cole, H.A., and Clark, D.J. (2014). RSC-dependent constructive and destructive interference between opposing arrays of phased nucleosomes in yeast. Genome research *24*, 1637-1649.

Gkikopoulos, T., Schofield, P., Singh, V., Pinskaya, M., Mellor, J., Smolle, M., Workman, J.L., Barton, G.J., and Owen-Hughes, T. (2011). A role for Snf2-related nucleosome-spacing enzymes in genome-wide nucleosome organization. Science (New York, NY) *333*, 1758-1760.

Gossett, A.J., and Lieb, J.D. (2012). In vivo effects of histone H3 depletion on nucleosome occupancy and position in Saccharomyces cerevisiae. PLoS genetics *8*, e1002771.

Gutin, J., Sadeh, R., Bodenheimer, N., Joseph-Strauss, D., Klein-Brill, A., Alajem, A., Ram, O., and Friedman, N. (2018). Fine-Resolution Mapping of TF Binding and Chromatin Interactions. Cell reports *22*, 2797-2807.

Hartley, P.D., and Madhani, H.D. (2009). Mechanisms that specify promoter nucleosome location and identity. Cell *137*, 445-458.

Hennig, B.P., Bendrin, K., Zhou, Y., and Fischer, T. (2012). Chd1 chromatin remodelers maintain nucleosome organization and repress cryptic transcription. EMBO reports *13*, 997-1003.

Ito, T., Bulger, M., Pazin, M.J., Kobayashi, R., and Kadonaga, J.T. (1997). ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. Cell *90*, 145-155.

Jaiswal, R., Choudhury, M., Zaman, S., Singh, S., Santosh, V., Bastia, D., and Escalante, C.R. (2016). Functional architecture of the Reb1-Ter complex of Schizosaccharomyces pombe. Proceedings of the National Academy of Sciences of the United States of America *113*, E2267-2276.

Knoll, K.R., Eustermann, S., Niebauer, V., Oberbeckmann, E., Stoehr, G., Schall, K., Tosi, A., Schwarz, M., Buchfellner, A., Korber, P., *et al.* (2018). The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling. Nature structural & molecular biology *25*, 823-832.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science (New York, NY) *184*, 868-871.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell *98*, 285-294.

Krietenstein, N., Wal, M., Watanabe, S., Park, B., Peterson, C.L., Pugh, B.F., and Korber, P. (2016). Genomic Nucleosome Organization Reconstituted with Pure Proteins. Cell *167*, 709-721.e712.

Kubik, S., Bruzzone, M.J., Challal, D., Dreos, R., Mattarocci, S., Bucher, P., Libri, D., and Shore, D. (2019). Opposing chromatin remodelers control transcription initiation frequency and start site selection. Nature structural & molecular biology *26*, 744-754.

Kubik, S., O'Duibhir, E., de Jonge, W.J., Mattarocci, S., Albert, B., Falcone, J.L., Bruzzone, M.J., Holstege, F.C.P., and Shore, D. (2018). Sequence-Directed Action of RSC Remodeler and General Regulatory Factors Modulates +1 Nucleosome Position to Facilitate Transcription. Molecular cell *71*, 89-102.e105.

Lai, W.K.M., and Pugh, B.F. (2017). Understanding nucleosome dynamics and their links to gene expression and DNA replication. Nature reviews Molecular cell biology *18*, 548-562.

Langst, G., Bonte, E.J., Corona, D.F., and Becker, P.B. (1999). Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. Cell *97*, 843-852.

Lieleg, C., Ketterer, P., Nuebler, J., Ludwigsen, J., Gerland, U., Dietz, H., Mueller-Planitz, F., and Korber, P. (2015). Nucleosome spacing generated by ISWI and CHD1 remodelers is constant regardless of nucleosome density. Molecular and cellular biology *35*, 1588-1605.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Lusser, A., Urwin, D.L., and Kadonaga, J.T. (2005). Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. Nature structural & molecular biology *12*, 160-166.

McKnight, J.N., Jenkins, K.R., Nodelman, I.M., Escobar, T., and Bowman, G.D. (2011). Extranucleosomal DNA binding directs nucleosome sliding by Chd1. Molecular and cellular biology *31*, 4746-4759.

McKnight, J.N., Tsukiyama, T., and Bowman, G.D. (2016). Sequence-targeted nucleosome sliding in vivo by a hybrid Chd1 chromatin remodeler. Genome research *26*, 693-704.

Ngo, H.B., Lovely, G.A., Phillips, R., and Chan, D.C. (2014). Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation. Nature communications *5*, 3077.

Ocampo, J., Chereji, R.V., Eriksson, P.R., and Clark, D.J. (2016). The ISW1 and CHD1 ATP-dependent chromatin remodelers compete to set nucleosome spacing in vivo. Nucleic acids research *44*, 4625-4635.

Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). Science (New York, NY) *183*, 330-332.

Olins, D.E., and Olins, A.L. (2003). Chromatin history: our view from the bridge. Nature reviews Molecular cell biology *4*, 809-814.

Parnell, T.J., Huff, J.T., and Cairns, B.R. (2008). RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. The EMBO journal *27*, 100-110.

Parnell, T.J., Schlichter, A., Wilson, B.G., and Cairns, B.R. (2015). The chromatin remodelers RSC and ISW1 display functional and chromatin-based promoter antagonism. eLife *4*, e06073.

Patel, A., McKnight, J.N., Genzor, P., and Bowman, G.D. (2011). Identification of residues in chromodomain helicase DNA-binding protein 1 (Chd1) required for coupling ATP hydrolysis to nucleosome sliding. The Journal of biological chemistry *286*, 43984-43993.

Pointner, J., Persson, J., Prasad, P., Norman-Axelsson, U., Stralfors, A., Khorosjutina, O., Krietenstein, N., Svensson, J.P., Ekwall, K., and Korber, P. (2012). CHD1 remodelers regulate nucleosome spacing in vitro and align nucleosomal arrays over gene coding regions in S. pombe. The EMBO journal *31*, 4388-4403.

Rawal, Y., Chereji, R.V., Qiu, H., Ananthakrishnan, S., Govind, C.K., Clark, D.J., and Hinnebusch, A.G. (2018). SWI/SNF and RSC cooperate to reposition and evict promoter nucleosomes at highly expressed genes in yeast. Genes & development *32*, 695-710.

Rhee, H.S., and Pugh, B.F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell *147*, 1408-1419.

Rippe, K., Schrader, A., Riede, P., Strohner, R., Lehmann, E., and Langst, G. (2007). DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes. Proceedings of the National Academy of Sciences of the United States of America *104*, 15635-15640.

Satchwell, S.C., Drew, H.R., and Travers, A.A. (1986). Sequence periodicities in chicken nucleosome core DNA. Journal of molecular biology *191*, 659-675.

Simic, R., Lindstrom, D.L., Tran, H.G., Roinick, K.L., Costa, P.J., Johnson, A.D., Hartzog, G.A., and Arndt, K.M. (2003). Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. The EMBO journal *22*, 1846-1856.

Smolle, M., Venkatesh, S., Gogol, M.M., Li, H., Zhang, Y., Florens, L., Washburn, M.P., and Workman, J.L. (2012). Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. Nature structural & molecular biology *19*, 884-892.

Stockdale, C., Flaus, A., Ferreira, H., and Owen-Hughes, T. (2006). Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes. The Journal of biological chemistry *281*, 16279-16288.

Thomas, J.O., and Furber, V. (1976). Yeast chromatin structure. FEBS letters *66*, 274-280.

Torigoe, S.E., Patel, A., Khuong, M.T., Bowman, G.D., and Kadonaga, J.T. (2013). ATP-dependent chromatin assembly is functionally distinct from chromatin remodeling. eLife *2*, e00863.

Tsankov, A., Yanagisawa, Y., Rhind, N., Regev, A., and Rando, O.J. (2011). Evolutionary divergence of intrinsic and trans-regulated nucleosome positioning sequences reveals plastic rules for chromatin organization. Genome research *21*, 1851-1862.

Tsukiyama, T., Palmer, J., Landel, C.C., Shiloach, J., and Wu, C. (1999). Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. Genes & development *13*, 686-697.

Udugama, M., Sabri, A., and Bartholomew, B. (2011). The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. Molecular and cellular biology *31*, 662-673.

Valouev, A., Johnson, S.M., Boyd, S.D., Smith, C.L., Fire, A.Z., and Sidow, A. (2011). Determinants of nucleosome organization in primary human cells. Nature *474*, 516-520.

van Bakel, H., Tsui, K., Gebbia, M., Mnaimneh, S., Hughes, T.R., and Nislow, C. (2013). A compendium of nucleosome and transcript profiles reveals determinants of chromatin architecture and transcription. PLoS genetics *9*, e1003479.

van Holde, K.E. (1989). Chromatin (New York: Springer).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

Varga-Weisz, P.D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P.B. (1997). Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. Nature *388*, 598-602.

Whitehouse, I., Rando, O.J., Delrow, J., and Tsukiyama, T. (2007). Chromatin remodelling at promoters suppresses antisense transcription. Nature *450*, 1031-1035.

Wiechens, N., Singh, V., Gkikopoulos, T., Schofield, P., Rocha, S., and Owen-Hughes, T. (2016). The Chromatin Remodelling Enzymes SNF2H and SNF2L Position Nucleosomes adjacent to CTCF and Other Transcription Factors. PLoS genetics *12*, e1005940.

Wippo, C.J., Israel, L., Watanabe, S., Hochheimer, A., Peterson, C.L., and Korber, P. (2011). The RSC chromatin remodelling enzyme has a unique role in directing the accurate positioning of nucleosomes. The EMBO journal *30*, 1277-1288.

Yamada, K., Frouws, T.D., Angst, B., Fitzgerald, D.J., DeLuca, C., Schimmele, K., Sargent, D.F., and Richmond, T.J. (2011). Structure and mechanism of the chromatin remodelling factor ISW1a. Nature *472*, 448-453.

Yan, C., Chen, H., and Bai, L. (2018). Systematic Study of Nucleosome-Displacing Factors in Budding Yeast. Molecular cell *71*, 294-305.e294.

Yang, J.G., Madrid, T.S., Sevastopoulos, E., and Narlikar, G.J. (2006). The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. Nature structural & molecular biology *13*, 1078-1083.

Yarragudi, A., Miyake, T., Li, R., and Morse, R.H. (2004). Comparison of ABF1 and RAP1 in chromatin opening and transactivator potentiation in the budding yeast Saccharomyces cerevisiae. Molecular and cellular biology *24*, 9152-9164.

Yen, K., Vinayachandran, V., Batta, K., Koerber, R.T., and Pugh, B.F. (2012). Genome-wide nucleosome specificity and directionality of chromatin remodelers. Cell *149*, 1461-1473.

Zhang, Y., Moqtaderi, Z., Rattner, B.P., Euskirchen, G., Snyder, M., Kadonaga, J.T., Liu, X.S., and Struhl, K. (2009). Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. Nature structural & molecular biology *16*, 847-852.

Zhang, Z., Wippo, C.J., Wal, M., Ward, E., Korber, P., and Pugh, B.F. (2011). A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. Science (New York, NY) *332*, 977-980.

Zhou, C.Y., Johnson, S.L., Lee, L.J., Longhurst, A.D., Beckwith, S.L., Johnson, M.J., Morrison, A.J., and Narlikar, G.J. (2018). The Yeast INO80 Complex Operates as a Tunable DNA Length-Sensitive Switch to Regulate Nucleosome Sliding. Molecular cell *69*, 677-688.e679.

Acknowledgments

We thank Kevin Schall, Nils Krietenstein and Maria Walker for providing purified proteins, Michael Wolff and Ulrich Gerland for helpful discussions of nucleosome positioning mechanisms, Stefan Krebs and Helmut Blum at the Laboratory for Functional Genome Analysis (LAFUGA, Gene Center, LMU München) for high throughput sequencing, and Sigurd Braun for access to and help with the plating robot. This study was funded by the German Research Foundation (SFB1064 to P.K.; SFB860, SPP1935, and EXC 2067/1-390729940 to P.C.), and the European Research Council (grant agreement No 693023 to P.C.).

Author contributions

Conceptualization: PK, SE, EO, KPH; Data curation: EO, VN; Formal analysis: EO, VN; Funding acquisition, Project administration, Supervision: PK, KPH, SE, PC, CLP; Investigation: EO, VN, SW, LF, MM, AS; Methodology: EO, VN, SE, SW, LF, PK; Validation: EO, VN, MM, AS, SW, LF, PK, SE; Visualization: EO, VN, PK, SW, LF; Writing original draft: PK, SE, EO; Writing – review & editing: PK, SE, EO, VN, KPH, LF, SW, CLP, PC.

Competing interests

The authors declare no competing interests.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

Methods

Organisms as source for materials used in experiments.
The pGP546 yeast genomic plasmid library was expanded from the clonal plates provided by Open Biosystems. For generation of genomic plasmid libraries, the *S. pombe* strain Hu0303 (Ekwall group) and *E. coli* strain (ATCC 11303 strain, 14380, Affymetrix) were used.
INO80 wild-type and mutant complexes, Chd1 and FACT were expressed in *Trichoplusiani insect cells. Spodoptera frugiperda sf21* insect cells were used for virus production. The *Saccharomyces cerevisiae* ISW1a and Fun30 remodelers were purified from the correspondingly TAP-tagged yeast strains, Ioc3-TAP, Fun30-Tap, as provided by Open Biosystems. Yeast ISW2 was purified from strain YTT480 (ISW2-2xFLAG, Tsukiyama et al., 1999). Reb1 was purified from *E. coli* BL21 (DE3) cd+ cells. The *Drosophila* embryo histones were prepared from the *Drosophila melanogaster* strain OregonR.

Embryonic *D. melanogaster* histones, whole-genome plasmid libraries and salt gradient dialysis

Embryonic *D. melanogaster* histone purification. The preparation of embryonic *D. melanogaster* histones octamers was carried out as described in Krietenstein et al. 2012 and Simon and Felsenfeld, 1979. Briefly, 50 g of 0-12 hours old *D. melanogaster* embryos were dechorionated in 3 % sodium hypochlorite, washed with dH_20 and resuspended in 40 mL lysis-buffer (15 mM K·HEPES pH 7.5, 10 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 0.2 mM PMSF, 10 % glycerol). Embryos were homogenized (Yamamoto homogenizer), filtered through cloth and centrifuged at 6,500 *g* for 15 min. Nuclei (brownish light pellet) were washed 3 times with 50 mL sucrose-buffer (15 mM K·HEPES pH 7.5, 10 mM KCl, 5 mM $MgCl_2$, 0.05 mM EDTA, 0.25 mM EGTA, 1 mM DTT, 0.2 mM PMSF, 1.2 % sucrose) and resuspended in 30 mL sucrose-buffer containing 3 mM $CaCl_2$. To obtain mononucleosomes, nuclei were incubated for 10 min at 26 °C with 6250 Units MNase (Sigma-Aldrich). Reaction was stopped with 10 mM EDTA, nuclei were pelleted and resuspended in 6 mL TE (10 mM Tris·HCl pH 7.6, 1 mM EDTA) containing 1 mM DTT and 0.2 mM PMSF followed by 30 to 45 min of rotation at 4 °C. Nuclei were centrifuged for 30 min at 15,300 *g* at 4 °C. Solubilized mononucleosomes are found in the supernatant, which was applied to a pre-equilibrated hydroxyapatite column. After washing the hydroxyapatite column with 0.63 M KCl, histone octamers were eluted with 2 M KCl, concentrated and stored in 50 % glycerol and 1x Complete (Roche) protease inhibitors without EDTA at -20 °C.

Whole-genome plasmid library expansion. The *S. cerevisiae* genomic plasmid library (pGP546) was originally described in Jones et al. 2008 and purchased as a clonal glycerol stock collection from Open Biosystems. Library expansion was carried out via a Singer ROTOR plating machine (Singer Instruments) (8-12 rounds, 3 replicas). After 16 hours, colonies were combined into 3x2 L of LB medium containing 50 µg/mL kanamycin and grown for 4 hours. Cells were harvested and subjected to Plasmid Giga Preparation (PC 10 000 Kit, Macherey&Nagel).
For *S. pombe* and *E. coli* plasmid library generation, genomic *S. pombe* (Hu0303) and *E. coli* (type B cells, ATCC 11303 strain, 14380, Affymetrix) DNA was fragmented by a limited SauIIIA or AluI digest. Fragmented DNA was ligated into pJET1.2 vector (ThermoFisher Scientific) and transformed into electrocompetent DH5α cells. Cells were plated on LB plates containing 100 µg/mL ampicillin, grown for 16 - 20 hours, combined in LB medium containing 100 µg/mL ampicillin and grown for another 4 hours. Plasmids was extracted with Plasmid Mega Preparation Kit (PC 2000 Kit, Macherey&Nagel).

Salt gradient dialysis (SGD). For low, medium and high assembly degrees, 10 µg of plasmid library DNA (*S. cerevisiae*, *S. pombe* or *E. coli*) was mixed with ~2, 4 or 8 µg of *Drosophila* embryo histone octamers, respectively, in 100 µl assembly buffer (10 mM Tris·HCl, pH 7.6, 2 M NaCl, 1 mM EDTA, 0.05 % IGEPAL CA630, 0.2 µg BSA). Samples were transferred to Slide-A-lyzer mini dialysis devices, which were placed in a 3 L beaker containing 300 mL of high salt buffer (10 mM Tris·HCl pH 7.6, 2 M NaCl, 1 mM EDTA, 0.05 % IGEPAL CA630, 14.3 mM β-mercaptoethanol), and dialyzed against a total of 3 L low salt buffer (10 mM Tris·HCl pH 7.6, 50 mM NaCl, 1 mM EDTA, 0.05 % IGEPAL CA630, 1.4 mM β-mercaptoethanol) added continuously via a peristaltic pump over a time course of 16 h while stirring. β-mercaptoethanol was added freshly to all buffers. After complete transfer of low salt buffer, samples were dialyzed against 1 L low salt buffer for 1 h at room temperature. DNA concentration of the SGD chromatin preparations was estimated with a DS-11+ spektrophotometer (Denovix) and could be stored at 4 °C for several weeks. To estimate the extent of the assembly degree, an aliquot of the sample was subjected to MNase digestion (as described below) for MNase-ladder read out.

Purifications of chromatin remodeling enzymes

Expression and purification of INO80 complex and respective mutants. Exact strategy for recombinant expression of *S. cerevisiae* INO80 complex in insect cells and complex purification is described in the accompanying paper Krietenstein et al. Briefly, MultiBac technology (Trowitzsch et al., 2010) was applied to generate two baculoviruses carrying coding sequences for *S. cerevisiae* Ino80 (2xFlag), Rvb1, Rvb2, Arp4, Arp5-His, Arp8, Actin, Taf14, Ies1, Ies2, Ies3, Ies4, Ies5, Ies6 and Nhp10 which were subcloned into pFBDM vectors and sequence verified by Sanger Sequencing (GATC Services at Eurofins Genomics). High Five (Hi5) insect cells (BTI-TN-5B1-4 Invitrogen) were co-infected with two or three baculoviruses 1/100 (v/v) each for expression purposes. The recombinantly expressed INO80 complex and respective INO80 mutant complexes were purified from insect cells according to (Tosi, Haas et al. 2013), which resulted in a pure and monodisperse sample. Shortly, cells were resuspended in lysis buffer (50 mM Tris·HCl pH 7.9, 500 mM NaCl, 10 % glycerol, 1 mM DTT, SIGMAFASTTM protease inhibitor cocktail), sonified (Branson Sonifier, 3x 20 s with 40 % duty cycle and output control 3-4) and cleared by centrifugation (Sorvall Evolution RC, SS34 rotor, 15,000 *g*). The supernatant was incubated with anti-Flag M2 Affinity Gel (Sigma-Aldrich) and centrifuged for 15 min at 1,000 *g* and 4 °C. The anti-Flag resin was washed with buffer A (25 mM K·HEPES pH 8.0, 500 mM KCl, 10 % glycerol, 0.025 mM IGEPAL CA630, 4 mM $MgCl_2$, 1 mM DTT) and buffer B (25 mM K·HEPES pH 8.0, 200 mM KCl, 10 % glycerol, 0.02 mM IGEPAL CA630, 4 mM $MgCl_2$, 1 mM DTT). Recombinant INO80 complex was eluted with buffer B containing 1.6 mg Flag Peptide (Sigma-Aldrich). Anion exchange chromatography (MonoQ 5/50 GL, GE Healthcare) was used for further purification which resulted in a monodisperse and clear INO80 complex. Using standard cloning techniques, three INO80(2xFlag) HSA domain mutants (HQ1, HQ2, HQ1/2, Figure 4D), one N-

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

terminal deletion mutant (Ino80$^{\Delta N}$, deletion of the first 461 amino acids of the N terminus of Ino80) and two INO80 (2xFlag) Nhp10 module mutants (ΔNhp10 (INO80 complex without Ies1, Ies3, Ies5 and Nhp10 but with Ino80 N-terminus) and HMGII (Figure 4G) were generated and integrated into baculoviruses using MultiBac Technology. Expression and purification of mutant INO80 complexes was carried out as described above. The INO80 core complex from *Chaetomium thermophilum* (equivalent to the *S. cerevisiae* N-terminal deletion mutant) was essentially purified as described in Eustermann et al., 2018.

Expression and purification of full-length Chd1 and FACT. Hi5 cells (600 mL) were grown in ESF-921 media (Expression Systems) and infected with V1 virus for full-length Chd1 (tagged with a N-terminal 6×His tag, followed by a MBP tag, and a tobacco etch virus protease cleavage site) or FACT (Spt16 carries an N-terminal 6×His tag, followed by an MBP tag, and a tobacco etch virus protease cleavage site) for protein expression. Cells were grown for 72 hours at 72 °C and subsequently harvested by centrifugation (238 *g*, 4 °C, 30 min). Supernatant was discarded and cell pellets resuspended in lysis buffer (300 mM NaCl, 20 mM Na·HEPES pH 7.4, 10 % (v/v) glycerol, 1 mM DTT, 30 mM imidazole pH 8.0, 0.284 µg/mL leupeptin, 1.37 µg/mL pepstatin A, 0.17 mg/mL PMSF, 0.33 mg/mL benzamidine). Resuspended cells were snap frozen and stored at -80 °C.

All protein purifications were performed at 4 °C. Frozen cell pellets were thawed and lysed by sonication. Lysates were cleared using centrifugation (18,000 *g*, 4 °C, 30 min and 235,000 *g*, 4 °C, 60 min). The supernatant containing Chd1 was filtered with 0.8-µm syringe filters (Millipore) and applied onto a GE HisTrap HP 5 mL (GE Healthcare). The column was washed with 10 column volumes (CV) lysis buffer, 5 CV high salt buffer (1 M NaCl, 20 mM Na·HEPES pH 7.4, 10 % (v/v) glycerol, 1 mM DTT, 30 mM imidazole pH 8.0, 0.284 µg/mL leupeptin, 1.37 µg/mL pepstatin A, 0.17 mg/mL PMSF, 0.33 mg/mL benzamidine), and 5 CV lysis buffer. Chd1 was eluted using a 40-minutes gradient of 0-100 % elution buffer (300 mM NaCl, 20 mM Na·HEPES pH 7.4, 10 % (v/v) glycerol, 1 mM DTT, 500 mM imidazole pH 8.0, 0.284 µg/mL leupeptin, 1.37 µg/mL pepstatin A, 0.17 mg/mL PMSF, 0.33 mg/mL benzamidine). Fractions containing Chd1 were pooled and subjected to dialysis/TEV protease digestion for 16 hours (300 mM NaCl, 20 mM Na·HEPES pH 7.4, 10 % (v/v) glycerol, 1 mM DTT, 30 mM imidazole with 2 mg His6-TEV protease).

The dialyzed sample was again applied to a GE HisTrap HP 5 mL. The flow-through, which contained cleaved tag-less Chd1, was concentrated using an Amicon Millipore 15 mL 50,000 MWCO centrifugal concentrator. The concentrate was applied to a GE S200 16/600 pg size exclusion column in 300 mM NaCl, 20 mM Na·HEPES pH 7.4, 10 % (v/v) glycerol, 1 mM DTT. Fractions containing Chd1 were concentrated to ~100 µM. The sample was aliquoted, snap frozen and stored at -80 °C.

FACT was purified as above with minor modifications. After dialysis, the sample was subjected to a tandem GE HisTrap HP 5 mL and GE HiTrap Q 5 mL columns combination. After sample application, the columns were washed with lysis buffer and the HisTrap removed. FACT was eluted by applying a high salt buffer gradient from 0-100 % high salt buffer (1 M NaCl, 20 mM Na·HEPES pH 7.4, 10 % (v/v) glycerol, 1 mM DTT, 30 mM imidazole pH 8.0). Fractions with FACT were applied to a GE S200 16/600 pg size exclusion column. Peak fractions with FACT were concentrated to a concentration of ~60 µM, aliquoted, snap frozen, and stored at -80 °C.

Expression and purifications of ISW1a, ISW2 and Fun30. Tandem affinity purification of ISW1a (TAP-Ioc3) and Fun30 (TAP-Fun30) was performed as follows: Cultures were grown in YPD media, harvested cells were washed once with water. The cells were lysed in buffer E (20 mM Na·HEPES pH 7.5, 350 mM NaCl, 10 % glycerol, 0.1 % Tween, and 0.5 mM DTT) and protease inhibitors by grinding in the presence of liquid nitrogen. Lysates were clarified at 40,000 *g* at 4 °C for 1 h. Cleared lysates were incubated with IgG-Sepharose (GE Healthcare) at 4 °C for 2 h and eluted by TEV protease (Invitrogen) cleavage at 4 °C overnight. The elutions were incubated with calmodulin affinity resin (Agilent Technology) in buffer E plus 2 mM $CaCl_2$ at 4 °C for 2 h and eluted in buffer E plus 10 mM EGTA.

ISW2 (FLAG-Isw2) was purified as follows: Cleared lysate was incubated with Anti-FLAG M2 affinity gel (Sigma-Aldrich) at 4 °C for 1 h and eluted with 0.1 mg/mL 3X FLAG peptide (Sigma-Aldrich). E-buffer (20 mM Na·HEPES pH 7.5, 350 mM NaCl, 10 % glycerol, 0.1 % Tween, and 0.5 mM DTT) was used during the entire purification.

Purified proteins were concentrated with VIVASPIN concentrators (Sartorius) and dialyzed against E-Buffer with 1 mM DTT. Subunit compositions were confirmed by SDS-PAGE (Figure S1A) and mass spectrometry.

Expression and purification of *S. cerevisiae* Reb1. Purification of *S. cerevisiae* Reb1 was essentially carried out as described in (Krietenstein et al., 2016). Briefly, using BY4741 genomic *S. cerevisiae* DNA the coding sequence for Reb1 was amplified by PCR and cloned into pET21b (Novagen) via InFusion cloning (Clontech) with a Streptavidin tag at the C terminus. Correct sequences were verified via Sanger sequencing (GATC Services at Eurofins Genomics). Expression plasmids were transformed into BL21 (DE3) cd^{+} cells. Three liters of LB medium supplemented with 600 mg/L ampicillin were inoculated with 200 mL pre-culture. Cells were grown at 37 °C to an OD_{600} of 0.6 (WPA CO8000 cell density meter). Induction was carried out by addition of IPTG to a final concentration of 1 mM. Cells were grown overnight at 18 °C, harvested by centrifugation (3,500 rpm, Sorvall Evolution RC) and stored at -80 °C. Cells were resuspended in lysis buffer (50 mM Tris·HCl pH 7.9, 500 mM NaCl, 7 % glycerol, 1 mM DTT, 7 % sucrose and protease inhibitor (SIGMAFAST™ protease inhibitor cocktail, 1:100), sonicated (Branson Sonifier 250, 5 min at 40-50 % duty cycle and output control 4) and cleared by centrifugation (Sorvall Evolution RC, SS34 rotor, 15,000 *g*). The supernatant was dialyzed over night against 2 L low salt buffer (25 mM K·HEPES pH 8.0, 150 mM KCl, 7 % glycerol, 4 mM $MgCl_2$, 1 mM DTT). Cation ion exchange chromatography (HiTrap SP HP 5 mL, elution buffer: 25 mM K·HEPES pH 8.0, 1 M KCl, 7 % glycerol, 4 mM $MgCl_2$, 1 mM DTT) followed by size exclusion chromatography (Superdex 200 10/300, buffer: 25 mM K·HEPES pH 8.0, 200 mM KCl, 7 % glycerol, 4 mM $MgCl_2$, 1 mM DTT) were used for purification. Peak fractions were analyzed by Coomassie SDS-PAGE. Fractions containing Reb1 were pooled, concentrated and stored at -80 °C.

Preparation of mononucleosomes with recombinant human octamers. Canonical human histones were provided by The Histone Source – Protein Expression and Purification (PEP) Facility at Colorado State University. Lyophilized individual human histones were resuspended in 7 M guanidinium chloride, mixed at a 1.2-fold molar excess of

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

H2A/H2B and dialyzed against 2 M NaCl for 16 h. Histone octamers were purified by size exclusion chromatography (HILoad 16/600 Superdex 200 column, GE Healthcare) and stored at -20 °C in 50 % glycerol.
We used fluorescein-labeled Widom 601 DNA (Lowary and Widom 1998) with 80 bp extranucleosomal DNA (80N0 orientation) harboring an in vivo ChIP-Exo verified Reb1 binding site (Rhee and Pugh 2012) of *S. cerevisiae* gene yGL167c (Reb1 binding motif: TTACCC) 64 or 84 bp distant to the 601 sequence. The DNA template (yGL267c_601 and yGL167c_20bp_601) was amplified via PCR, purified by anion exchange chromatography (HiTrap DEAE FF, GE Healthcare) and vacuum concentrated. DNA and assembled histone octamer were mixed in 1.1-fold molar excess of DNA at 2 M NaCl. Over a time-period of 17 h at 4 °C the NaCl concentration was reduced to a final concentration of 50 mM NaCl. Again, anion exchange chromatography was used to purify reconstituted nucleosome core particle (NCP) which were then dialyzed to 50 mM NaCl. NCPs were concentrated to 1 mg/mL and stored at 4 °C.

ATPase Assay. As described previously (Eustermann et al., 2018; Knoll et al., 2018), we applied an NADH-based ATPase assay (Kiianitsa et al., 2003) to determine INO80's ATPase rate. 15 nM INO80 were incubated at 30 °C in a final volume of 50 µl assay buffer (25 mM K·HEPES pH 8.0, 50 mM KCl, 5 mM $MgCl_2$, 0.1 mg/mL BSA) with 0.5 mM phosphoenolpyruvate, 2 mM ATP, 0.2 mM NADH and 25 units/mL lactate dehydrogenase/pyruvate kinase (Sigma-Aldrich) to monitor the NADH dependent fluorescence signal in non-binding, black, 384-well plates (Greiner) at an excitation wavelength of 340 nm and an emission wavelength of 460 nm over a 40-min period. We used the Tecan Infinite M1000 (Tecan) plate reader for read out. For all samples, ATPase activity was determined at maximum INO80 WT ATPase activity. ATPase activity was stimulated with 50 nM, 25 nM and 12.5 nM Reb1 site-80N0 mononucleosomes with or without WT Reb1 at indicated concentrations. Using maximal initial linear rates corrected for the buffer blank, we calculated final ATP turnover rates.

Genome-wide remodeling reaction. All remodeling reactions, except Chd1-containing reactions, were performed at 30 °C in 100 µL with final buffer conditions of 26.6 mM Na·HEPES pH 7.5, 1 mM Tris·HCl pH 7.6, 85.5 mM NaCl, 8 mM KCl, 10 mM ammonium sulfate, 10 mM creatine phosphate (Sigma-Aldrich), 3 mM $MgCl_2$, 2.5 mM ATP, 0.1 mM EDTA, 0.6 mM EGTA, 1 mM DTT, 14 % glycerol, 20 ng/µl creatine kinase (Roche Applied Science).
Chd1-containing reactions were performed in 26.6 mM Na·HEPES pH 7.5, 1 mM Tris·HCl pH 7.6, 50 mM NaCl, 10 mM creatine phosphate (Sigma-Aldrich), 3 mM $MgCl_2$, 2.5 mM ATP, 0.1 mM EDTA, 0.6 mM EGTA, 1 mM DTT, 14 % glycerol, 20 ng/µl creatine kinase. If called for, 10 nM of remodeling enzyme (but 50 nM Chd1/FACT), 40 nM Reb1 and 20 Units of BamHI (NEB) was added. Before full-length Chd1 (in high-salt buffer) was added to the reaction, it was diluted together with FACT into low salt buffer. For that, full-length Chd1 and purified FACT was mixed in a 1.2:1 molar ratio in high salt buffer (300 mM NaCl, 20 mM Na·HEPES pH 7.4, 10 % (v/v) glycerol, 1 mM DTT), incubated on ice for 5 min and then diluted to 30 mM NaCl final concentration.
Remodeling reactions were started by adding 10 µl SGD chromatin corresponding to about 1 µg DNA assembled into nucleosomes and terminated by adding 0.8 Units apyrase (NEB) followed by incubation at 30 °C for 30 min.

MNase-seq. After apyrase addition, remodeling reactions were supplemented with $CaCl_2$ to a final concentration of 1.5 mM and digested with 100 Units MNase to generate mostly monoucleosomal DNA. Chd1-reaction were incubated with 20 Units MNase to get the same extent of mononucleosomal DNA. 10 mM EDTA and 0.5 % SDS (final concentrations) were added to stop the MNase digest. After proteinase K treatment for 30 min at 37 °C, samples were ethanol precipitated and electrophoresed for 1.5 - 2 h at 100 V using a 1.5 % agarose gel in 1x Tris-acetate-EDTA (TAE) buffer. Mononucleosome bands were excised and purified with PureLink Quick Gel Extraction Kit (ThermoFisher Scientific).
For library preparation, 10-50 ng of mononucleosomal DNA was incubated with 1.25 Units Taq polymerase (NEB), 3.75 Units T4 DNA polymerase (NEB) and 12.5 Units T4-PNK (NEB) in 1x ligation buffer (B0202S, NEB) for 15 min at 12 °C, 15 min at 37 °C and 20 min at 72 °C. To ligate NEBNext Adaptors (0.75 µM final concentration, NEBNext Multiplex Oligos Kit) to the DNA, samples were incubated with T4 DNA ligase (NEB) at 25 °C for 15 min, followed by incubation with 2 Units USER enzyme (NEB) for 10 min at 37 °C. Fragments were purified using 2 volumes AMPure XP beads (Beckman Coulter) and amplified for 8-10 cycles using NEBNext Multiplex Oligos, Phusion High-Fidelity DNA Polymerase (1 U, NEB), deoxynucleotide solution mix (dNTP, 2.5 mM, NEB) and Phusion HF Buffer (1x, NEB). The following protocol was applied for amplification: 98 °C for 30 s, 98 °C for 10 s, 65 °C for 30 s, 72 °C for 30 s with a final amplification step at 72 °C for 5 min. DNA content was assessed by using Qubit dsDNA HS Assay Kit (Invitrogen). PCR reactions were applied to an 1.5 % agarose gel, needed fragment length (~270 bp) was excised and purified via PureLink Quick Gel Extraction Kit (ThermoFisher Scientific). DNA was measured again with Qubit dsDNA HS Assay Kit and diluted to a final concentration of 10 nM (calculation based on the assumption that the DNA fragment length is 272 bp, i. e., 147 bp nucleosomal DNA and 122 bp sequencing adaptor). Diluted samples were pooled according to sequencing reads (~6 Mio reads/ sample). The final pool was quantified with BioAnalyzer (Agilent) and analyzed on an Illumina HiSeq 1500 in 50 bp single-end mode (Laboratory for Functional Genome Analysis, LAFUGA, LMU Munich).

Data Processing. Sequencing data was mapped to the SacCer3 (R64), EF2 or *E. coli* strain B (REL606) genome using Bowtie (Langmead et al., 2009). Multiple matches were omitted. After mapping, data was imported into R Studio using GenomicAlignments (Lawrence et al., 2013). Every read was shifted by 73 bp to cover the nucleosome dyad and extended to 50 bp. Genome coverage was calculated and either aligned to in vivo +1 nucleosome positions (Xu et al., 2009), BamHI cut sites, Reb1 SLIM-ChIP hits (Gutin et al., 2018) or Reb1 PWM hits (Badis et al., 2008). Signal was normalized per gene in a 2001 bp window centered on the alignment point.
Heatmaps were sorted either by NFR length (distance between in vivo +1 and -1 nucleosome annotated by calling nucleosomes of *in vivo* MNase-seq by Tirosh) or by Reb1 binding score. For the latter, Reb1 SLIM-ChIP data (GSM2916407) was aligned to in vivo +1 nucleosome positions and sorted by signal strength in a 120 bp-window 160 bp upstream of every +1 nucleosome.
For promotor grouping according to Reb1 site orientation, Reb1 SLIM-ChIP hits which contain a PWM site (± 50 bp) and which are located within 400 bp upstream of in vivo +1 nucleosomes were used. Cluster 1 contains promotors where

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

the Reb1 PWM motif is located on the sense strand and cluster 2, where the Reb1 PWM motif is located on the antisense strand. Cluster 3 contains Reb1 sites at bidirectional promotors.

Data Resources

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE140614.

References

Badis, G., Chan, E.T., van Bakel, H., Pena-Castillo, L., Tillo, D., Tsui, K., Carlson, C.D., Gossett, A.J., Hasinoff, M.J., Warren, C.L., et al. (2008). A Library of Yeast Transcription Factor Motifs Reveals a Widespread Function for Rsc3 in Targeting Nucleosome Exclusion at Promoters. Molecular Cell *32*, 878–887.

Eustermann, S., Schall, K., Kostrewa, D., Lakomek, K., Strauss, M., Moldt, M., and Hopfner, K.-P. (2018). Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. Nature *556*, 386–390.

Gutin, J., Sadeh, R., Bodenheimer, N., Joseph-Strauss, D., Klein-Brill, A., Alajem, A., Ram, O., and Friedman, N. (2018). Fine-Resolution Mapping of TF Binding and Chromatin Interactions. Cell Reports *22*, 2797–2807.

Jones, G.M., Stalker, J., Humphray, S., West, A., Cox, T., Rogers, J., Dunham, I., and Prelich, G. (2008). A systematic library for comprehensive overexpression screens in Saccharomyces cerevisiae. Nat Meth *5*, 239–241.

Kiianitsa, K., Solinger, J.A., and Heyer, W.-D. (2003). NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. Anal. Biochem. *321*, 266–271.

Knoll, K.R., Eustermann, S., Niebauer, V., Oberbeckmann, E., Stoehr, G., Schall, K., Tosi, A., Schwarz, M., Buchfellner, A., Korber, P., et al. (2018). The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling. Nat Struct Mol Biol *25*, 823–832.

Krietenstein, N., Wippo, C.J., Lieleg, C., and Korber, P. (2012). Genome-Wide In Vitro Reconstitution of Yeast Chromatin with In Vivo-Like Nucleosome Positioning. In Methods in Enzymology, (Elsevier), pp. 205–232.

Krietenstein, N., Wal, M., Watanabe, S., Park, B., Peterson, C.L., Pugh, B.F., and Korber, P. (2016). Genomic Nucleosome Organization Reconstituted with Pure Proteins. Cell *167*, 709-721.e12.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology *10*, R25.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for Computing and Annotating Genomic Ranges. PLOS Computational Biology *9*, e1003118.

Simon, R.H., and Felsenfeld, G. (1979). A new procedure for purifying histone pairs H2A + H2B and H3 + H4 from chromatin using hydroxylapatite. Nucleic Acids Res *6*, 689–696.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *457*, 1033–1037.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figures

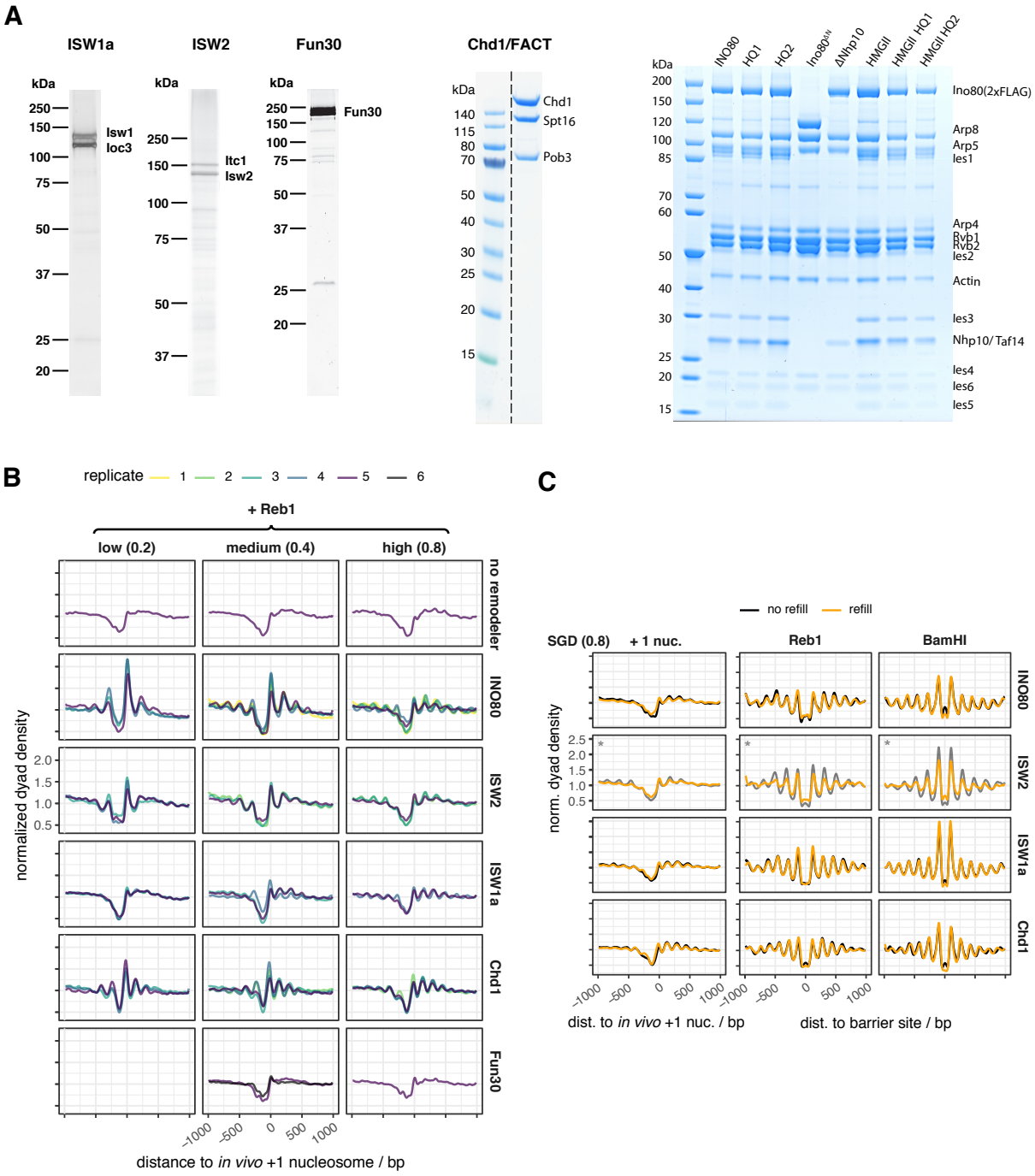


Figure S1, associated with Figures 1 and 2. (A) SDS-PAGE analyses of purified remodeler complexes. **(B)** Composite plots as in Figure 1D for individual replicates and the indicated combinations of remodeler, Reb1 and nucleosome density. "no remodeler" denotes absence of remodeler. **(C)** Composite plots aligned at *in vivo* +1 nucleosome positions (left), Reb1 (middle) or BamHI (right) sites for MNase-seq analysis of SGD chromatin assembled at high nucleosome density and incubated with the indicated remodelers as in Figure 2A (no refill) or with doubling remodeler concentration for the second half of incubation time (refill).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

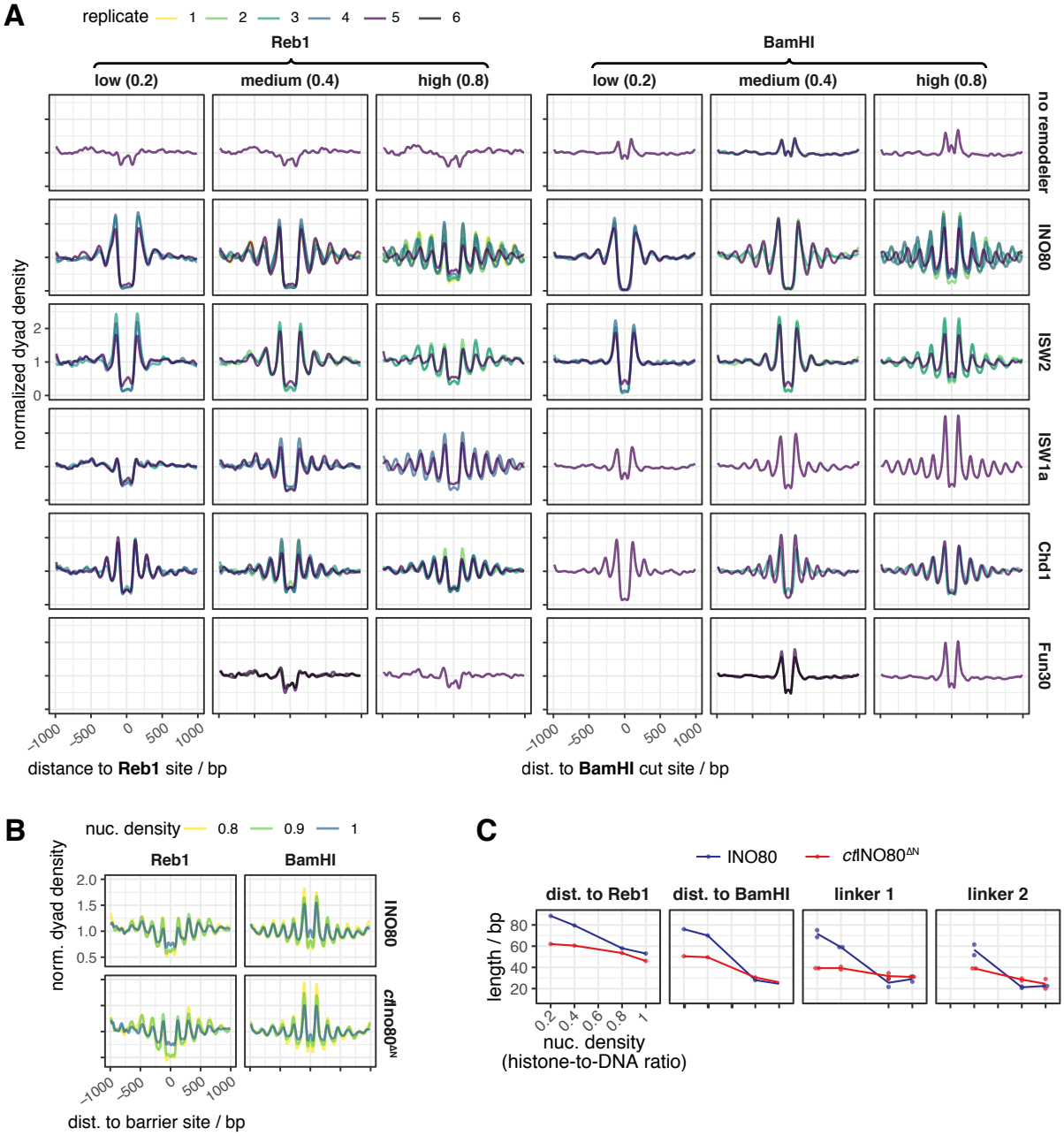


Figure S2, associated with Figure 2. (A) As Figure 2A but for individual replicates and indicated combinations of barrier, remodeler and nucleosome density. "no remodeler" denotes absence of remodeler. **(B)** As panel A, but for the indicated nucleosome densities and *S. cerevisiae* WT INO80 complex versus *C. thermophilum* core INO80 complex (ctINO80$^{\Delta N}$). **(C)** As Figure 2E, but for the indicated remodelers (as in panel B) and nucleosome densities.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

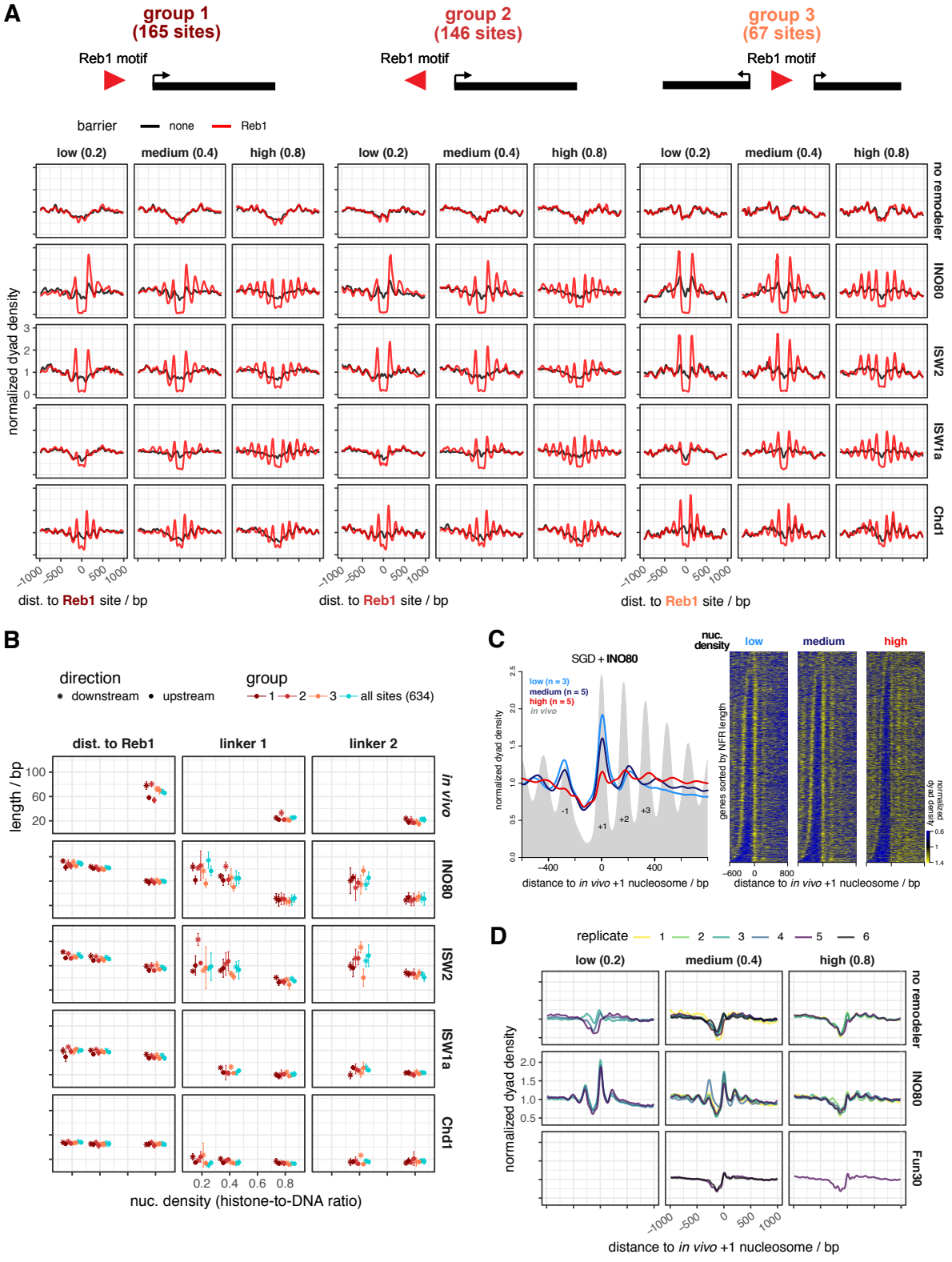


Figure S3, associated with Figure 2. (A) As Figure 2A but for indicated combinations of remodeler and nucleosome density. "no remodeler" denotes absence of remodeler. "no remodeler" denotes absence of Reb1. Tracks show merged data from replicates shown in Figures S1B, S2A. Reb1 sites were sorted into groups 1 to 3 (number of sites per group indicated) according to orientation of Reb1 sites relative to one or two genes. **(B)** As Figure 2C, but averages and standard deviation of values for up- and downstream lengths are shown for same groups as in panel A and for all anti-Reb1 SLIM-ChIP sites that contain a Reb1 PWM (light blue). **(C)** Composite plots (left) and heat maps (right) of MNase-seq analysis of *in vivo* chromatin or SGD chromatin reconstituted with the indicated nucleosome density and incubated with recombinant WT INO80 complex. Data are aligned at *in vivo* +1 nucleosome positions and heat maps are sorted from top to bottom by increasing NFR length. Traces with indicated replicate number (n) represent merged data for all replicates. Positions of -1, +1, +2, +3 nucleosomes of the *in vivo* pattern are labeled. **(D)** Composite plots as in panel C, right, but for individual replicates and the indicated remodelers. "no remodeler" denotes absence of remodeler.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

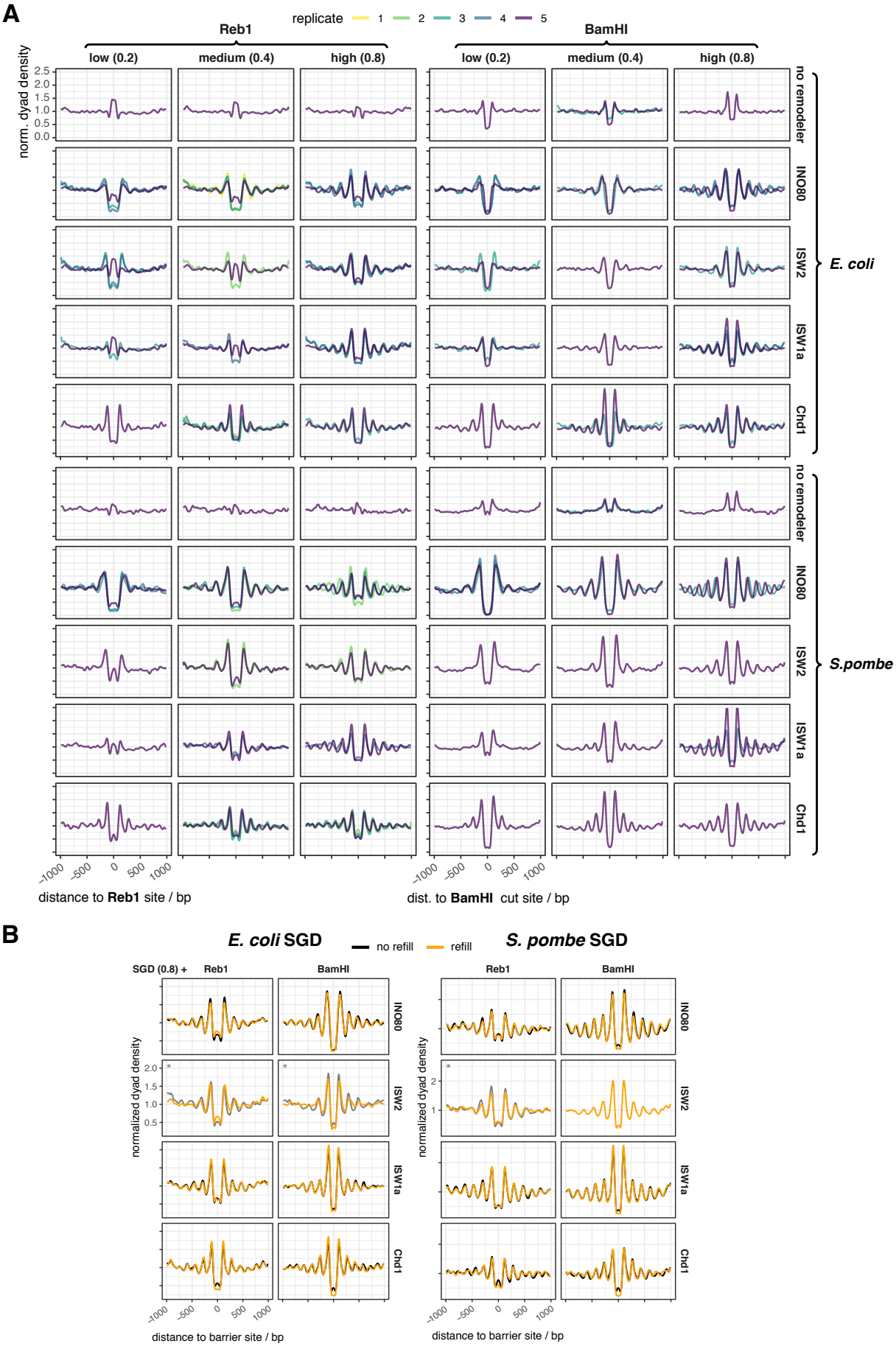


Figure S4, associated with Figure 3. (A) As Figure S2A, but for the indicated genomes. Replicate 5 corresponds to the reconstitution with mixed genomes. **(B)** As Figure S2B, but for the indicated genomes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.28.969618. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

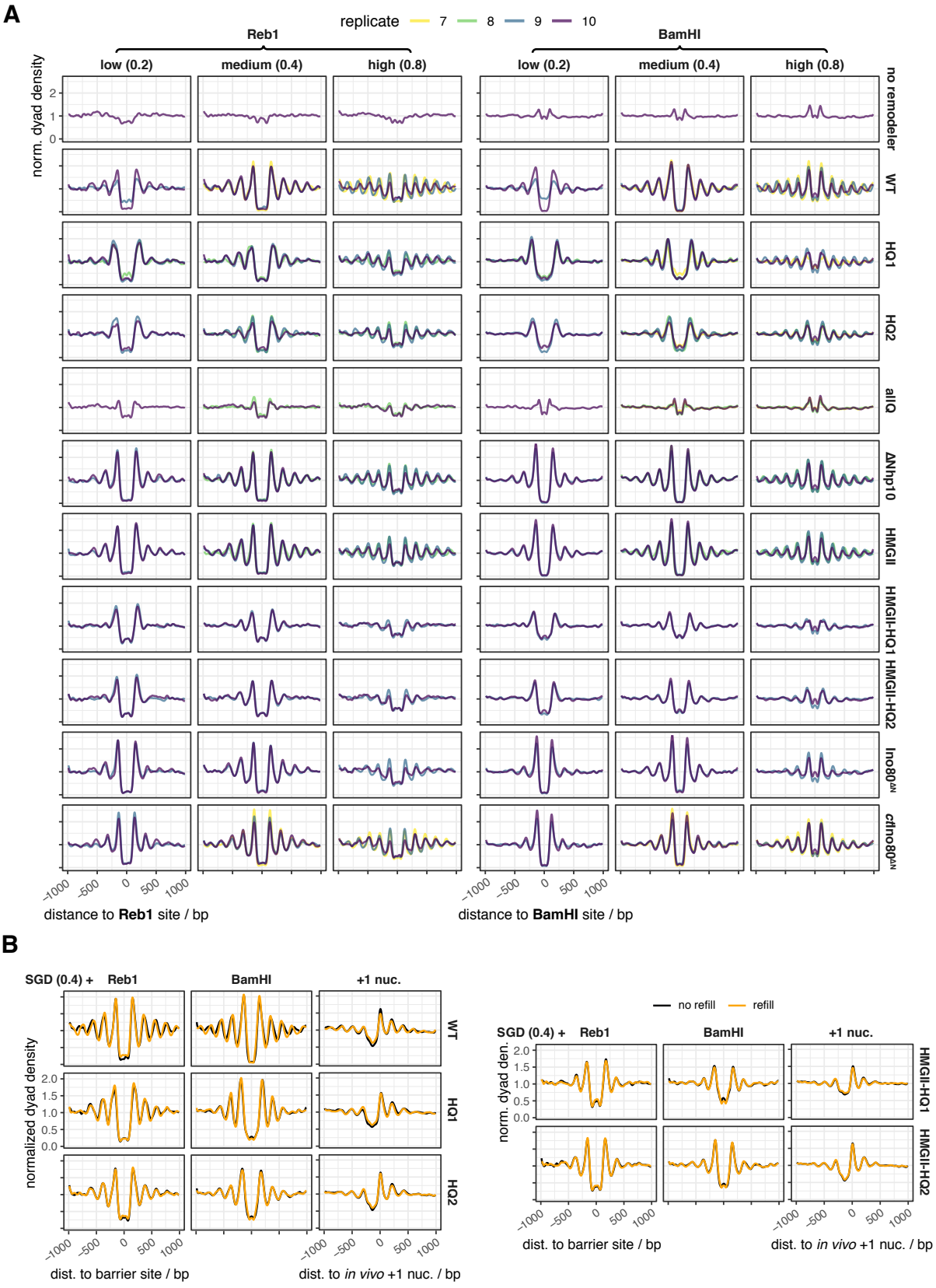


Figure S5, associated with Figures 5, 6. (A) As Figure S2A, but for the indicated WT and mutant INO80 remodelers and the *C. thermophilum* INO80 core complex (ctINO80$^{\Delta N}$). **(B)** As Figure S2B, but for indicated remodelers.

ACKNOWLEDGEMENTS

Zuallererst möchte ich meinem Doktorvater, PD Dr. Philipp Korber, danken für die großartigen Projekte, die vielen kreativen Ideen und Diskussionen, die immerwährende Unterstützung und das offene Ohr während meiner ganzen Promotionszeit. Ich weiß sehr zu schätzen, dass mir mehr als nur ein Projekt anvertraut wurde und ich immer in meinen doch sehr ehrgeizigen Zielsetzungen so selbstlos unterstützt wurde. Es war eine unglaublich spannende und lehrreiche Zeit, und ich hoffe etwas von dem Korber-*Spirit* verinnerlicht zu haben.

Des Weiteren danke ich meinem *Thesis Advisory Committee* und gleichzeitigen Kollaborateuren Dr. Tobias Straub und Dr. Sebastian Eustermann für ihre klare und ehrliche Bewertung, und hilfreichen Kommentare im TAC. Außerdem wäre es mir ohne Tobias nicht möglich gewesen, das wunderbare Gefühl einer selbst durchgeführten und gleichzeitig erfolgreichen Datenanalyse zu erleben. Und ohne Sebastians lebhafte Erklärungen wäre mein Verständnis seiner Strukturen um einiges kleiner.

Dann möchte ich meinen zahlreichen Kollaborateuren danken, ohne die der Blickwinkel auf meine Arbeit deutlich einseitiger geblieben wäre: zunächst Prof. Dr. Ulrich Gerland, Dr. Mark Heron und Michael Wolff für ihre bioinformatischen Analysen, und besonders Michael für die aufschlussreichen Diskussionen am Telefon. Auch möchte ich allen danken, die mich so unermüdlich mit hochreinem Protein(komplexen) versorgt haben: Dr. Shinya Watanabe und Prof. Dr. Craig Peterson, Dr. Lucas Farnung und Prof. Dr. Patrick Cramer, und allen voran das Hopfner-Labor mit Dr. Sebastian Eustermann, Dr. Kilian Knoll, Kevin Schall, Vanessa Niebauer und natürlich Prof. Dr. Karl-Peter Hopfner. Besonders hervorheben möchte ich Vanessa, mein erster *Padawan* in Sachen In-vitro Rekonstitution. Dann möchte ich Prof. Helmut Blum und Dr. Stefan Krebs danken für ihre Unterstützung in jeglichen Sequenzierangelegenheiten. Des Weiteren möchte ich Andrea Schmid danken, meiner dritten Hand im Labor, und Dr. Nils Krietenstein, der die Grundlage für fast all meine Projekte gelegt hat, und der mich die hohe Kunst der In-vitro Rekonstitution gelehrt hat.

Auch möchte ich allen danken die mir meine Promotionszeit mit Koch- und Kaffeepausen versüßt haben. Allen voran das alte „Northlab“ (Corinna, Nils, Lisa, Sarah, Maria), und Edith und Irene, die mir erst gezeigt haben, dass kochen in der Mittagspause möglich und auch sinnvoll ist. Dann dem neuen Korber-Labor (besonders Iris und Erika) für unendlichen Tee- und Kaffeenachschub in stressigen Zeiten gepaart mit aufmunternden Diskussionen, und der Bialetti-Crew, für zahlreiche, viel zu starke, dennoch wirksame Espressi.

Zu guter Letzt möchte ich Prof. Peter Becker danken, der eine wunderbare wissenschaftlich anregende und zugleich sehr angenehme Atmosphäre in seiner Abteilung geschaffen hat, die all dies erst ermöglicht hat.