Molecular basis of CENP-C association with the CENP-A nucleosome

at yeast centromeres

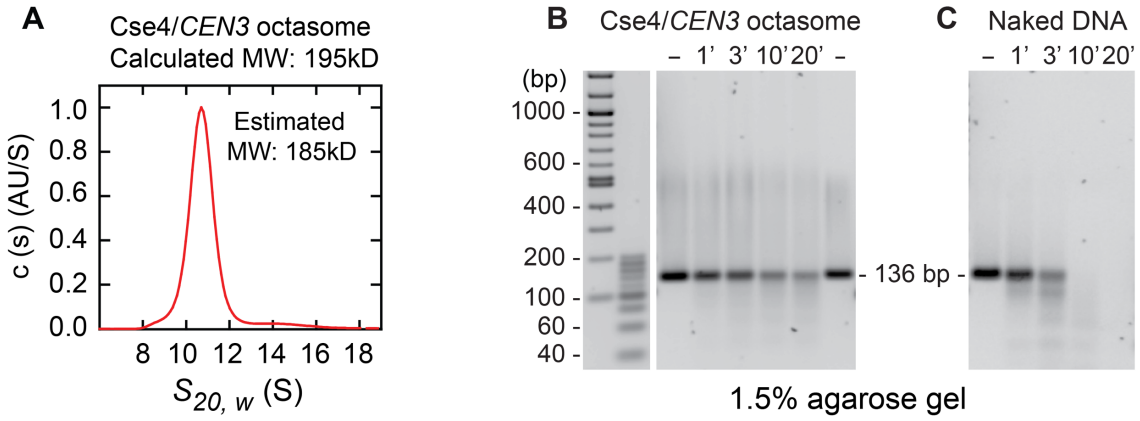
Hua Xiao, Feng Wang, Jan Wisniewski, Alexey K. Shaytan, Rodolfo Ghirlando, Peter C. FitzGerald, Yingzi Huang, Debbie Wei, Shipeng Li, David Landsman, Anna R. Panchenko, and Carl Wu

Supplemental Materials

Supplemental_Figures S1 to S11_and_Legends

Supplemental_Tables S1 and S3_and Legends

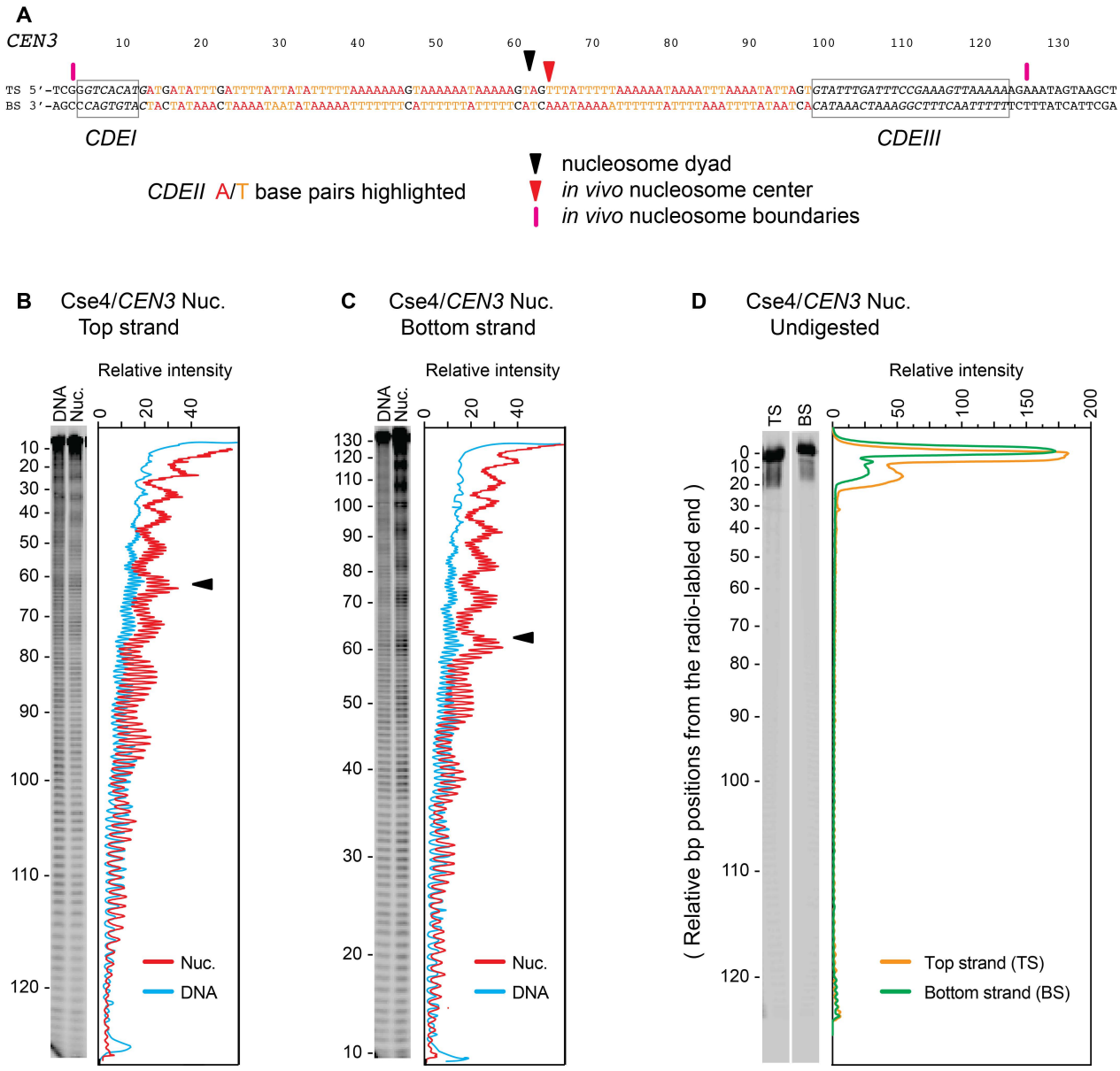
Supplementary Figures and Legends



Supplemental Figure S1 (Related to Figs 1 to 5). Analytical Ultracentrifugation (AUC) and MNase digestion of Cse4/*CEN3* nucleosomes.

(A) AUC of Cse4/*CEN3* nucleosome reconstituted on 130bp *CEN3* DNA (see Materials and Methods for details). Normalized absorbance *c(s)* distributions obtained for Cse4/*CEN3* nucleosome (red) at 340nM. The Cse4/*CEN3* nucleosome yields a sedimentation coefficient of 10.69 S (estimated molar mass 185 kDa).

(B, C) MNase digestion of reconstituted Cse4/*CEN3* nucleosomes. Samples were digested with 0.5 U of MNase for the indicated time (min). Digested samples were treated to 1% SDS, and DNA analyzed by EMSA and SYBR Green staining. Note that the 136bp DNA as indicated has 130bp *CEN3* DNA with additional 6bp from the restriction enzyme AvaI site (see Materials and Methods).

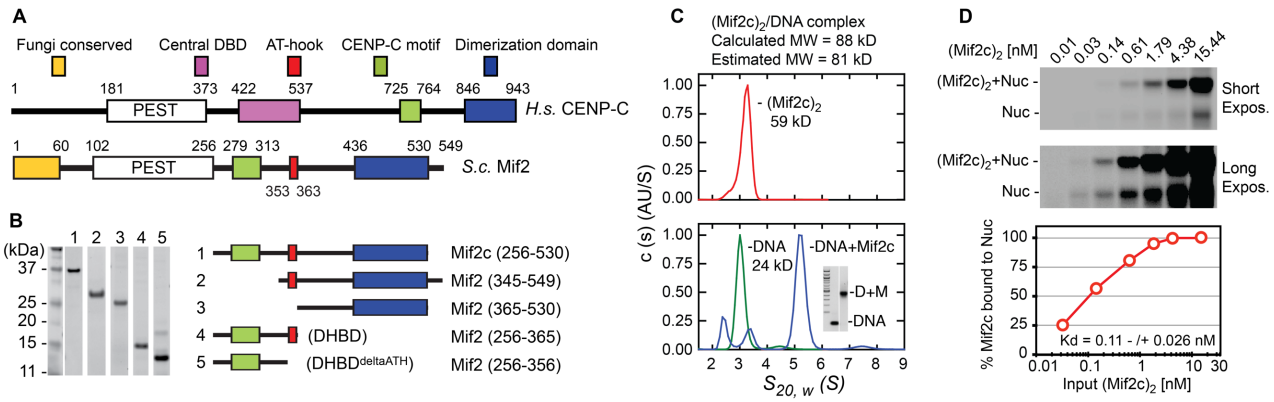


Supplemental Figure S2 (related to Figs 1 and 5). Hydroxyl radical footprinting of *CEN3* DNA and Cse4/*CEN3* nucleosomes.

(A) *CEN3* DNA sequence with the nucleosome dyad (Shaytan et al. 2017) and *in vivo* nucleosome center and boundaries indicated (Cole et al. 2011).

(B and C) Autoradiograms and densitometry scans of top and bottom DNA strands of ^{33}p-labeled *CEN3* DNA and *in vitro* reconstituted Cse4/*CEN3* nucleosomes. Cyan lines represent free DNA, and red lines nucleosomes. ^{33}p-labeled *CEN3* DNA and Cse4/*CEN3* nucleosome were partially digested with Iron(II)-EDTA (see Materials and Methods), DNA extracted and analyzed on a sequencing gel.

(D) Autoradiograms and densitometry scans of top (TS) and bottom (BS) strands of DNA extracted from undigested Cse4/*CEN3* nucleosome. Orange line represents top strand, and green line bottom strand.



Supplemental Figure S3 (Related to Figs. 1 to 6). Characterization of Mif2 core (Mif2c) binding to DNA and to Cse4/*CEN3* nucleosomes.

(A) Schematic representation of conserved domains of human CENP-C and the budding yeast homolog Mif2 (Brown, 1995; Cohen et al., 2008). Conserved structural features include the CENP-C motif and the dimerization domain. The N-terminal region of 60 amino acids (yellow) is conserved only among fungal proteins, and deletion of this region has no apparent growth phenotype (Cohen et al., 2008), but a recent study showed that Ame1–Okp1 directly associates with this region of Mif2 to facilitate outer kinetochore assembly (Hornung et al., 2014). Another notable difference is the presence of a minor groove-binding AT-hook motif in yeast Mif2 but absent from human CENP-C, which harbors a central DNA-binding domain (purple box) (Politi et al., 2002; Trazzi et al., 2009; Yang et al., 1996) containing a CENP-C-like motif (Kato et al., 2013), in addition to the classic CENP-C motif (green box). The PEST domain contains a high percentage of the four amino acid residues P, E, S, and T, and is involved in the regulation of protein stability (Lanini and McKeon, 1995). Sequences surrounding and including the CENP-C motif, sequences between the CENP-C motif and the dimerization domain, and the dimerization domain itself are all essential for protein function *in vivo* (Cohen et al., 2008). Thus, Mif2c

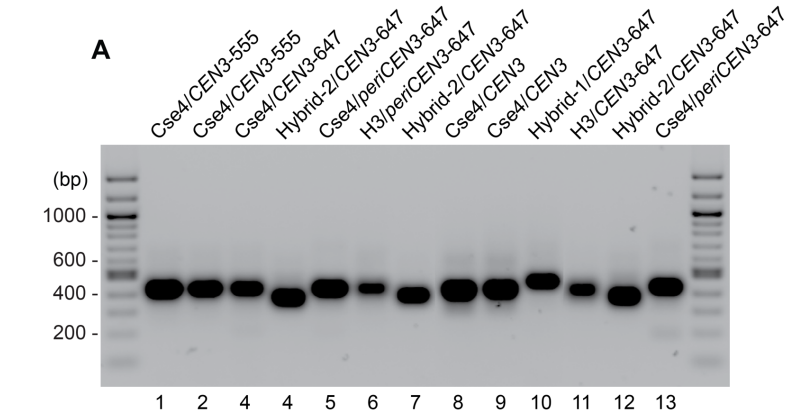
appears to contain all essential functions.

(B) Expression and purification of recombinant Mif2 derivatives (see Materials and Methods).

(C) Analytical ultracentrifugation demonstrates that Mif2c is a dimer in solution and when bound to DNA. The upper graph shows the absorbance *c(s)* of the free Mif2c indicating a single dimeric species (red line) with a sedimentation coefficient of 3.04 S and average molar mass of 59 ± 3 kDa (expected monomer mass is 32.047 kDa). To study the Mif2c/DNA complex formed with a 40bp dsDNA, the dsDNA was first characterized by sedimentation velocity and shown to primarily consist of single 3.08 S species with a molar mass of 24.5 kDa (green line, expected mass is 24.575 kDa). A 500µl sample containing Mif2c and the 40bp dsDNA at ~2µM concentration was subjected to analytical ultracentrifugation (the inset in the lower graph represents an agarose gel showing the free DNA and DNA/Mif2c complex, D+M). The major species observed at 5.23 S has an estimated molar mass of 81 kDa, indicating that a single Mif2c dimer binds to the dsDNA (blue line, calculated mass of 88.669 kDa). Mif2c binds to a Cse4/*CEN3* nucleosome also as a dimer (see Fig. 1C and D).

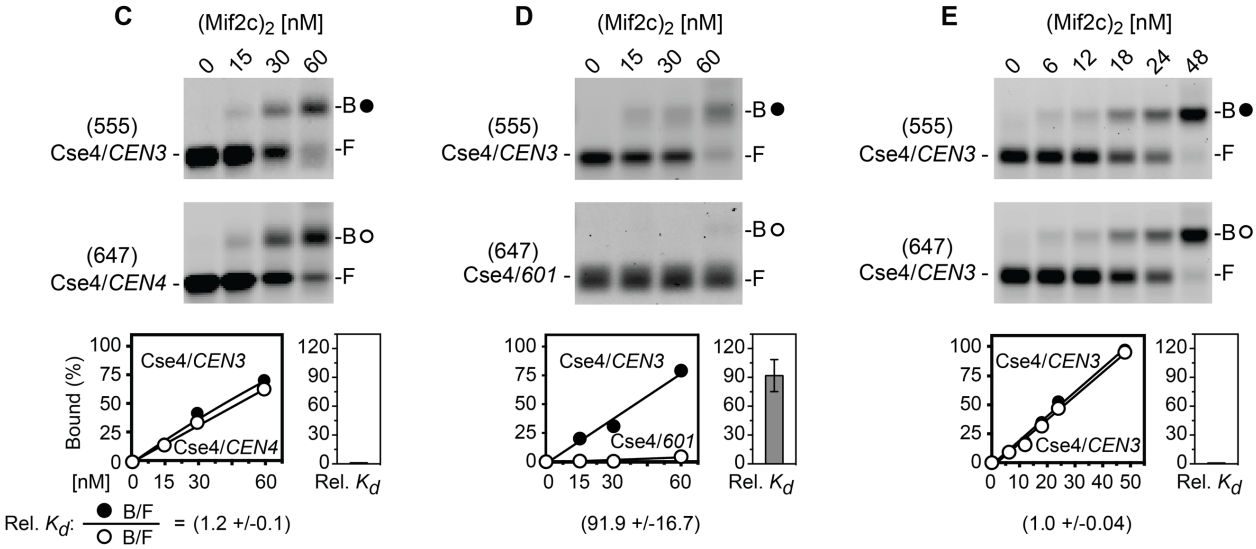
(D) Measurement of Mif2c binding affinity for the Cse4/*CEN3* nucleosome. To determine the dissociation constant, 0.39pmol of purified Mif2c was added to 0.46pmol of Cse4/*CEN3* nucleosome labeled with Alexa Fluor-647 in binding buffer containing 100mM NaCl and 0.4mg/ml BSA, and the volume was adjusted to 25µl, giving a final concentration of 15.44nM of Mif2c and 18.4nM of nucleosome, and incubated for 30 minutes at room temperature. Under these conditions, 100% of Mif2c was bound to nucleosomes. Then ~3-fold serial dilutions were made in the same binding buffer, and

the reactions were incubated for a further 50 minutes. Free and Mif2c bound nucleosomes were resolved on a 1.3% native agarose gel. The gel was scanned on a Typhoon scanner. The two top panels show two different levels of exposure of the same gel. Data was quantified using ImageQuant (Amersham Biosciences), and exported to EXCEL (Microsoft) and plotted as input Mif2c concentrations on log scale against % Mif2c bound to nucleosomes.



B

Ref Band	Volume	Relative concentration	For EMSA µl / reaction
1	50278365	2.0000000000	2.0
2	39529923	2.5438129378	2.5
3	34627933	2.9039195393	2.9
4	42355091	2.3741356610	2.4
5	46136907	2.1795290915	2.2
6	24450143	4.1127257351	4.1
7	33485205	3.0030196721	3.0
8	56777102	1.7710789603	1.8
9	52846879	1.9027941338	1.9
10	36154858	2.7812785484	2.8
11	26071277	3.8569929916	3.9
12	43188320	2.3283316402	2.3
13	41998945	2.3942680078	2.4



Supplemental Figure S4 (Related to Figs 2 and 3). Reconstitution of Cse4 and H3 nucleosomes and EMSA assays of their interactions with purified $(Mif2c)_2$.

(A) A representative scan of reconstituted nucleosomes analyzed on a 1.3% native agarose gel stained with SYBR green I.

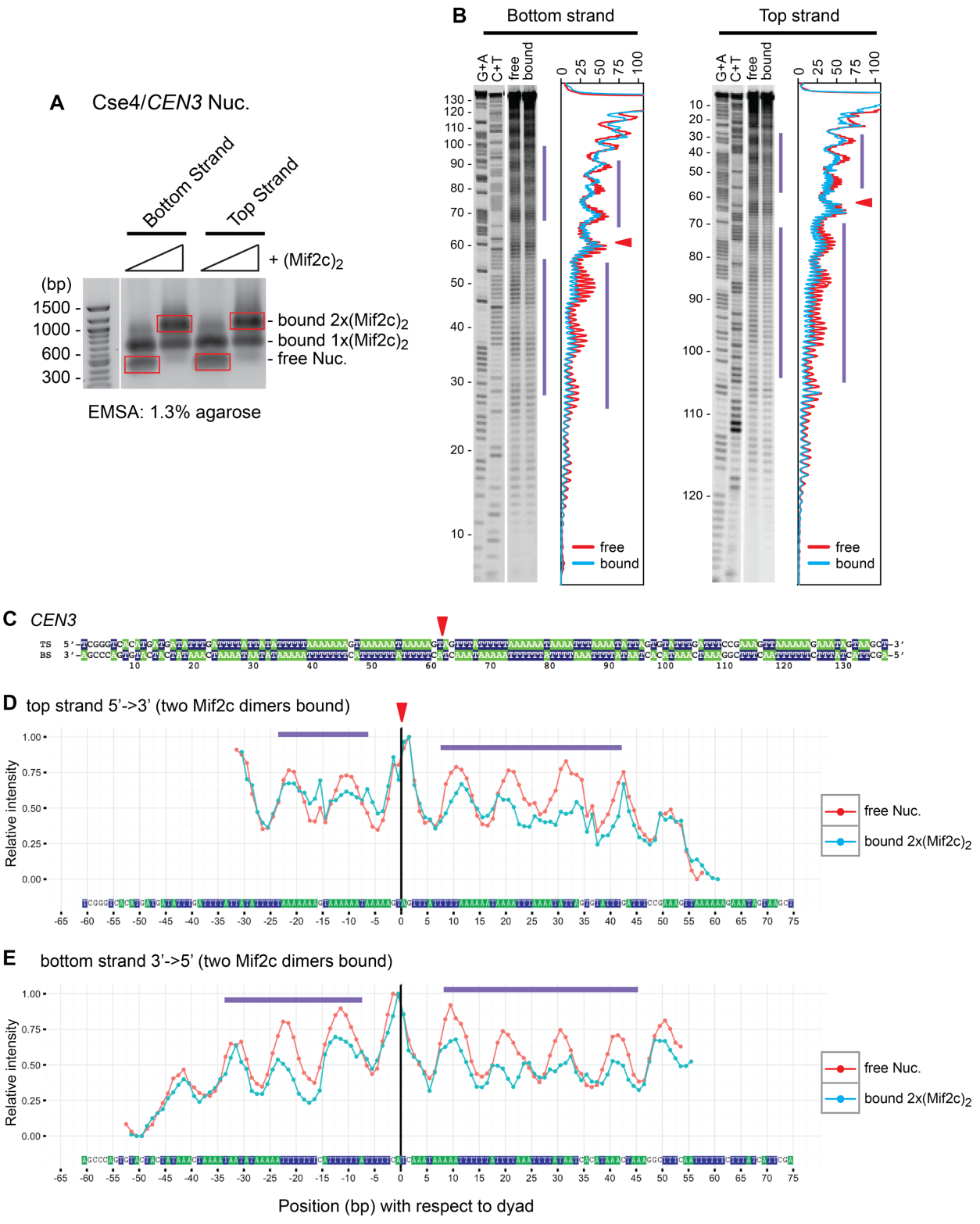
(B) Sample quantification of gel scan in panel A to determine the relative amounts (µl) of nucleosomes to be used in competitive EMSA assays.

(C) Mif2c binds *CEN3* and *CEN4* nucleosomes with similar affinities.

(D) Mif2c binds with strong preference to Cse4/*CEN3* nucleosome over Cse4/*601* nucleosome.

(E) Mif2c binds with very similar affinities to Alexa Flour-555 (top gel; closed circles indicate Mif2c-bound) and Alexa Flour-647 (bottom gel; open circles indicate Mif2c-bound) labeled Cse4/*CEN3* nucleosomes, demonstrating that the two fluorescent labels do not differentially affect the binding activities of Mif2c.

Competitive binding reactions, agarose gel electrophoresis, data collection, analysis and calculations were performed as in Figure 2.



Supplemental Figure S5 (related to Figs 4 and 5). Hydroxyl radical footprinting two

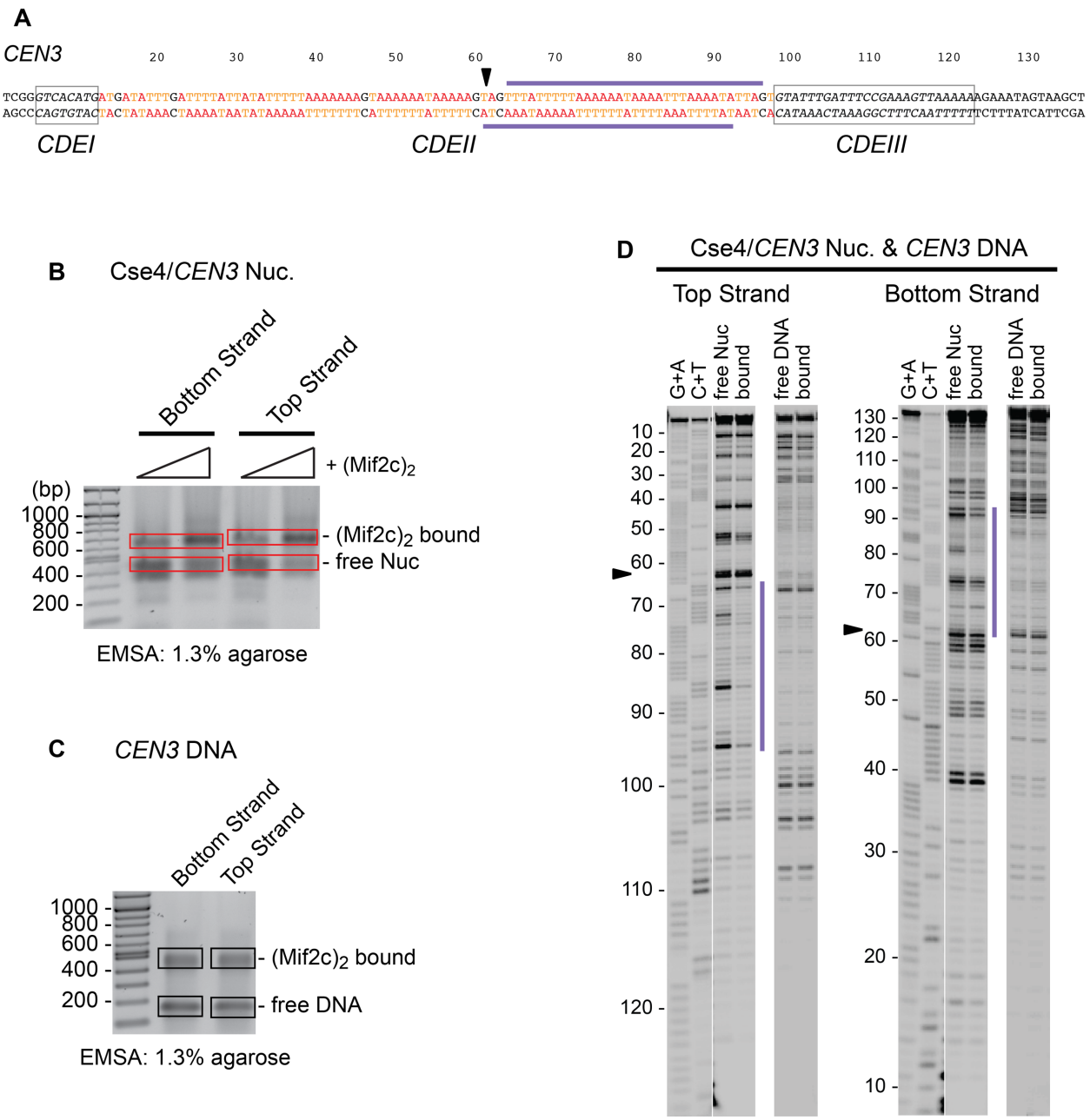
Mif2c dimers on Cse4/CEN3 nucleosome.

(A) Nucleosomes reconstituted on ^{33}p-labeled *CEN3* DNA were incubated with purified Mif2c, partially digested with Iron(II)-EDTA (see Materials and Methods), free and Mif2c-bound nucleosomes resolved on an agarose gel. Bands containing free nucleosomes and Mif2c-bound nucleosomes with two Mif2c dimers were excised, DNA extracted and analyzed on a sequencing gel.

(B) Autoradiograms and densitometry scans of top and bottom DNA strands of hydroxyl radical digested Cse4/*CEN3* nucleosome. Mobility markers were generated from A+G and C+T sequencing reactions of the same ^{33}p-labeled DNA fragments. Red lines represent free nucleosomes, and cyan lines Mif2c-bound nucleosomes.

(C) Position of the Cse4/*CEN3* nucleosome dyad symmetry axis on the *CEN3* DNA as identified by symmetry of hydroxyl-radical footprinting patterns (see Shaytan et al. 2017). Top and bottom strands are shown in base paired representation. Red arrow indicates dyad at position 61.5, about 2bp from the *in vivo* nucleosome center (Cole et al. 2011; see Supplemental Fig. 4C).

(D and E) Cleavage intensity profiles of DNA strands at specific nucleotides in free Cse4/*CEN3* nucleosomes and nucleosomes with two Mif2c dimers-bound. Intensity values of every profile are normalized from 0 to 1. Top strand plots are shown in 5’ to 3’ direction, while bottom strand plots in 3’ to 5’ direction. Blue bars indicate the span protected from hydroxyl radical cleavage conferred by binding of the Mif2c dimers to the nucleosome.

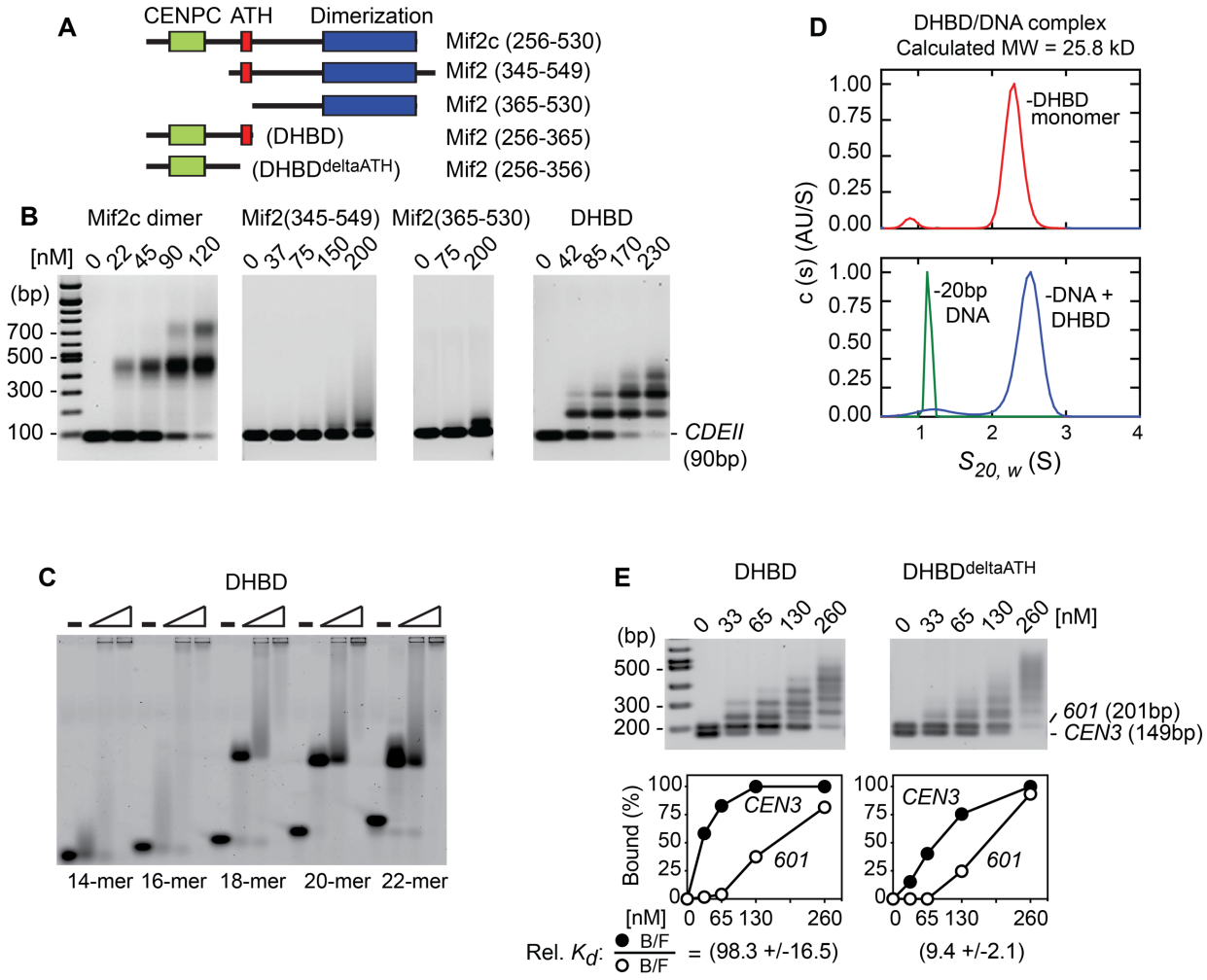


Supplemental Figure S6 (Related to Figs 4 and 5). DNase I footprinting to map Mif2c binding sites on Cse4/*CEN3* nucleosome and free *CEN3* DNA.

(A) DNA sequence of *CEN3* where A and T were colored in red and orange, respectively. Triangle indicates dyad of the Cse4/*CEN3* nucleosome. Purple bars indicate span of protection by Mif2c on the Cse4/*CEN3* nucleosome.

(B and C) Nucleosomes and DNA were incubated with purified Mif2c, and partially digested with DNase I. Bound and free nucleosomes and DNA were separated on a native agarose gel. Bands containing bound and free fractions (boxed) were excised from gel, DNA was recovered from gel slices and separated on a sequencing gel.

(D) Representative footprinting gels of top and bottom strands with mobility markers generated from A+G and C+T sequencing reactions of the same ^{33}P-labeled DNA fragments.



Supplemental Figure S7 (Related to Fig. 6). Analysis of DNA-binding activities of Mif2.

(A) Schematic representation of Mif2 expression constructs (see Fig. S1A and B).

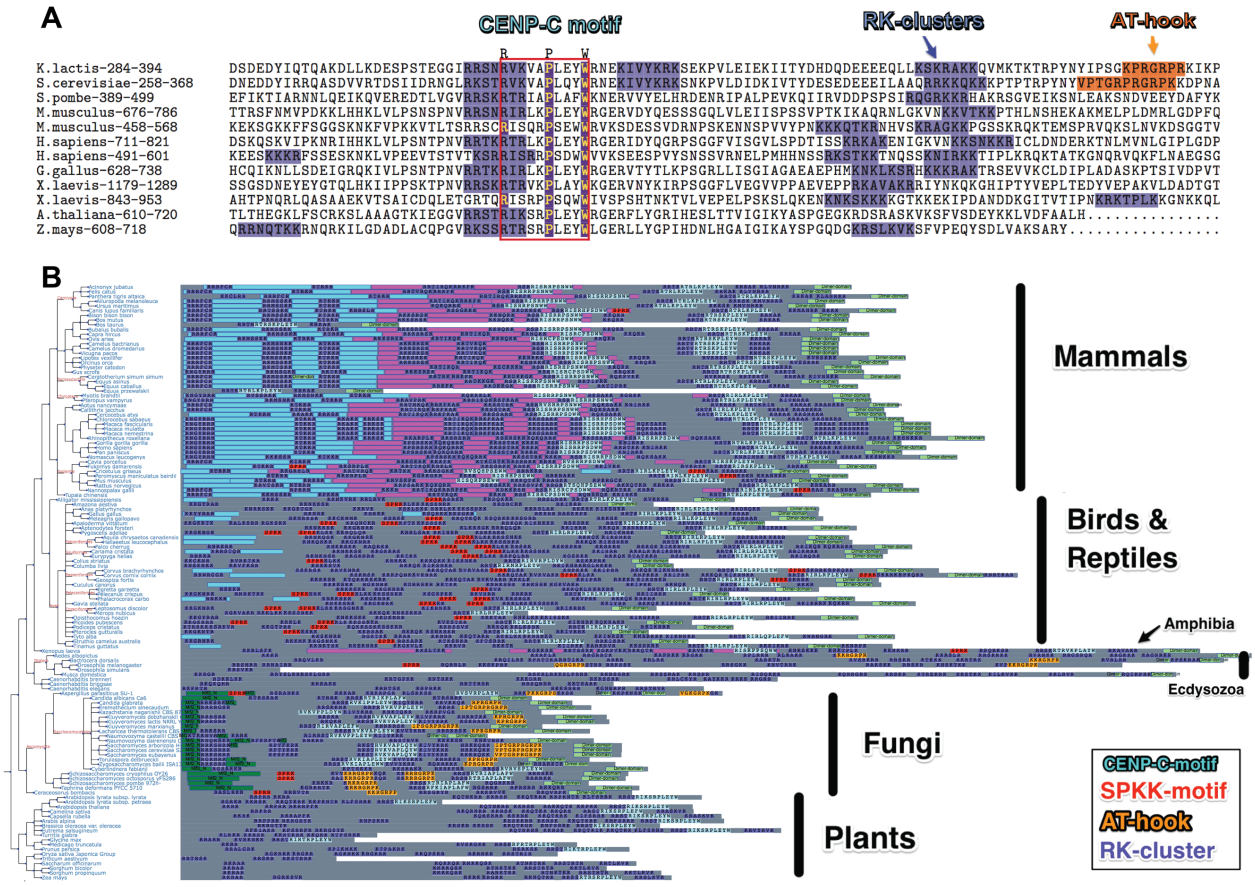
(B) EMSA assays of Mif2 and derivatives for their DNA binding activities. Purified Mif2 and derivatives (Fig. S1B) were added to DNA-binding reactions containing a 90bp *CDEII* DNA from *CEN3*, and the reactions were analyzed on a 1.3% native agarose gel. The DNA binding activity of Mif2 was located to amino acids 256-365 including the conserved CENP-C motif and an AT-hook motif. The stepwise mobility shifts on the 90bp DNA with

increasing protein concentration reflect association of multiple Mif2c dimers or DHBD monomers.

(C) EMSA assays using double stranded DNA oligos to determine the minimal site for binding of Mif2 DHBD.

(D) Analytical ultracentrifugation of Mif2(256-365) and its complex with a 20bp double stranded DNA shows that it is a monomer in solution and when bound to DNA.

(E) Mif2(256-356) lacking AT-hook retains modest selectivity for *CEN* DNA over *601* DNA. Equal amounts of 201bp *601* DNA (open circles) and 149bp *CEN3* DNA (closed circles) were mixed and incubated with increasing concentrations of purified Mif2 derivatives. After electrophoresis, the gel was stained with SYBR Green I and scanned. Data collection, analysis and calculations were performed as in Figure 2. Relative affinities are given in parenthesis below the graph. It should be noted that the amounts of bound complexes were inferred indirectly from the depletion of the distinct free DNA fragments upon Mif2 binding, since bound complexes of the two different DNA fragments could not be differentiated.

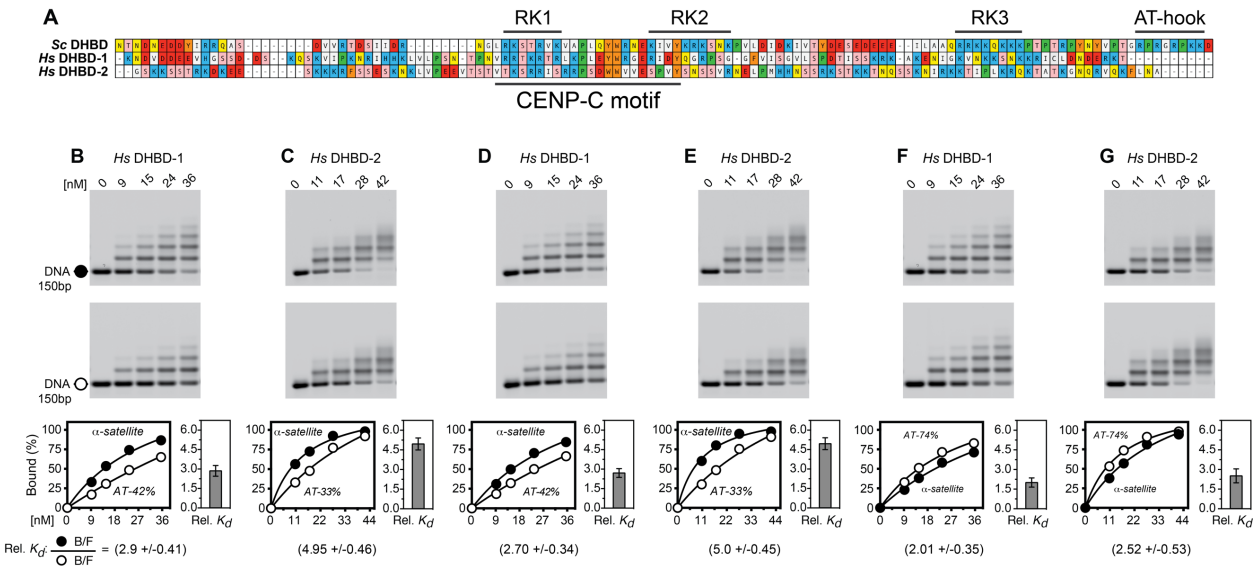


Supplemental Figure S8 (Related to Figure 7 and Supplemental Figure S9, S10). Analysis of sequence motifs and RK-clusters in CENP-C proteins.

(A) Location of DNA minor groove binding motifs and RK-clusters (= or >4 of 7 residues) in the vicinity of CENP-C-motif for a selected number of species. All protein sequences are aligned by CENP-C-motif.

(B) Birds-eye view of location and distribution of DNA minor groove binding motifs and RK-clusters in CENP-C proteins across different species grouped by phylogenetic kinship. Apart from AT-hooks, SPKK-motifs and RK-clusters, following protein domain are shown as identified by PFAM signatures: CENP-C_N in cyan, CENP-C_mid in magenta,

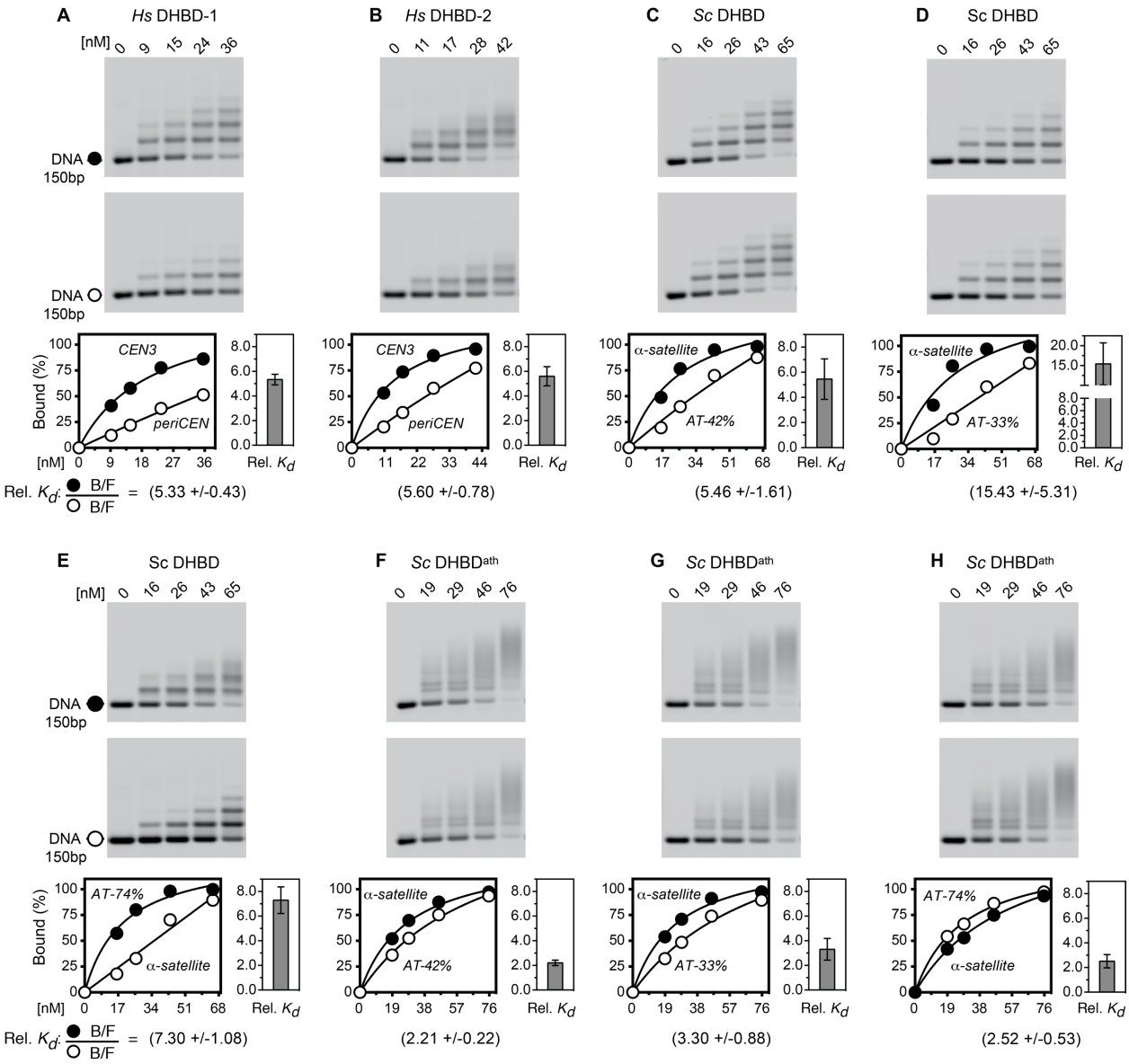
Dimerization domain (CENP-C_C) in light green. If more than one CENP-C-motif is present in a sequence both regions are presented.



Supplemental Figure S9 (Related to Supplemental Figures S10 and S11). Human CENP-C DNA-histone-binding domains (*Hs*DHBDs) bind preferentially to centromere DNAs.

(A) Sequence alignments of budding yeast Mif2 domain *Sc*DHBD (aa256-365) and human CENP-C domains *Hs*DHBD-1 (aa482-597) and *Hs*DHBD-2 (aa698-808). Blue: RK, Red: DE, Orange: FWY, Green: P, Pink: ST.

(B-G) EMSA was conducted with DHBDs and the indicated mixtures of DNAs. *AT-33%* and *AT-42%* are derived from the human NPM3 promoter and intron regions, respectively, and *AT-74%* from the mar del(10) neocentromere (see Supplemental Figure S11 for more details). Data collection and quantification were performed as in Figure 2.



Supplemental Figure S10 (Related to Figure 7 and Supplemental Figures S9 and S11). Gel shift analysis to compare DNA binding between budding yeast and human DHBDs.

Both yeast and human DHBDs binds preferentially to sequences of higher AT contents such as human α-satellite (58% A+T) and *AT-74%* (from the mar del(10) neocentromere) over *AT-33%* (from NPM3 promoter) and *AT-42%* (from NPM3 intron) sequences, and to

yeast *CEN* DNA (86% A+T) over *periCEN* DNA (65% A+T) (see Supplemental Figure S11 for more details). Competitive binding reactions, agarose gel electrophoresis, data collection, analysis and calculations were performed as in Figure 2.

A

Yeast DNA fragments

CEN3 GTCACATGATGATATTTGATTTTATTATATTTTTAAAAAAAGTAAAAAATAAAAAGTAGTTTATTTTTAAAAAATAAAATTTAAAATATTAGTGTATTTGATTTCCGAAAGTTAAAAAAGAAATAGTAAG (86%AT)

CEN10 TTAATCACGTGTTAAATAATTAATTTACTTTAAAATTTATTTTTTAATATAAAATATTTATTCTTTTTATTTAAAAATAAAAAACACAAAAAAACAATGTTTATGATTTCCGAACCTAAATATACACTTTACGA (85%AT)

periCEN3 GTATAAGGATTACTTGTCTTCTTTGCTACATATTGCTACCACTTCTATTACACAATAGTTTCAATAGCTTGCAGCGTAGCTAAACTCTAAAATTTATCTAAATCACTCATATAAACCGAACCCTTCCCCTTCT (65%AT)

Human DNA fragments

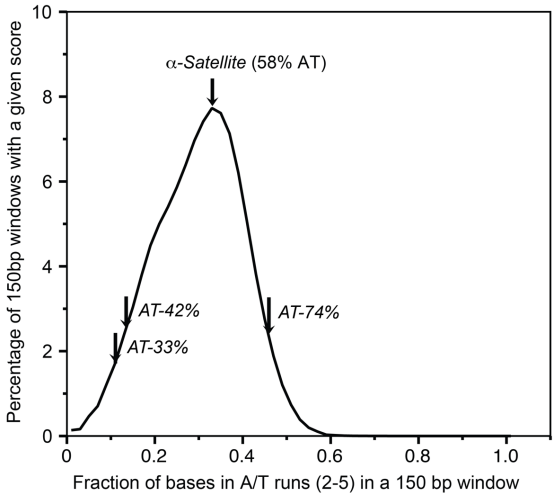
α-satellite CAGACCGCATTGAGGCCTTCGTTGGAAACGGGATTTCTTCATTTCATGCTAGACAGAAGAATTCTCAGTAACTTCTTTGTGCTGTGTGTATTCAACTCACAGAGTGGAACGTCCCTTTGCACAGAGCAGATTTGAAACACTCTTTTTGTA (58%AT)

AT-33 AGAGGGGGCCTCGCACGCACAGCCAAAGGGGGTGGCGAGATCCCTGCCCCAGATGTGTTGGGCGGGGCCCACGGCAGAGCGTCCCAGGTGTTCTGAGAGAAGAGACGGAGGTGGACGTGTAAATGACCCGTCCGCCCAAGGTTACACGTG (*NPM3 promoter*, 33%AT)

AT-42 TCTCCTGCGGGAACAGCGAAGCAGCCCTCCGCCCACACCCACGCCTTGAAACCCTCCGCATTCCCGCCTCACATCCCTACCTCTTACCGCCCCAGTACCACCCTCAGCCTCTCCCTTCACTAATACCGAAGAAAAAACTGTCCATAGTGA (*NPM3 intron*, 42%AT)

AT-74 TACCATAATGTTGTGTTGAATTTTATATTCTTAGAAAATCATCTGTCAAGGTGTTAACTAATGGCAAAGCATTTAATAAATCAGCATTCATGTATTCAGGTGCTCTGAATTATCTGACTTTTAAATTCTTACTTTATAAATGAGAAAATT (*neoCEN*, 74%AT)

B



Supplemental Figure S11 (Related to Figures 1 to 7; Supplemental Figures S1 to S7 and S9, S10). Enrichment of short, non-alternating A•T tracts in centromere DNA.

The sources of DNA sequences and their AT content (%) are indicated (see also Supplemental Figures S9 and S10).

(A) Centromere and non-centromere DNA sequences. A and T base pairs are highlighted in red and green, respectively.

(B) A bioinformatics analysis of A•T tracts 2-5bp in length over a 150bp window.

Supplementary Tables

	K. lactis	*S. cerevisiae*	*S. pombe*	*M. musculus*	*H. sapiens*	*G. gallus*	*X. laevis*	*C. elegans*	*D. melano-gaster*	*A. thaliana*	*Z. mays*
AT-hook	99.4%	99.9%	99.5%	0%	0%	0%	0%	0%	99.9%	0%	0%
SPKK	0%	0%	0%	0%	0%	0%	97%	0%	96%	0%	0%
RK	98.0%	97.4%	93.9%	97.0%	99.1%	98.3%	99.8%	97.1%	99.3%	99.3%	99.1%
CENP-C composition RK/ED	76/ 121	77/ 113	108/ 90	155/ 108	162/ 127	145/ 111	211/ 188	138/ 138	209/ 198	105/ 104	98/ 109

Supplemental Table S1 (Related to Fig. 7 and Supplemental Figs S8 and S9). Enrichment of DNA minor groove binding motifs and RK-clusters (= or >4 of 7 residues) in CENP-C proteins with respect to whole proteome for a given organism.

To characterize enrichment of a given motif, first the distribution of proteins with respect to the number of motif occurrences per protein was build, next, the position of CENP-C in this distribution was characterized by the percentile value, which is reported in every cell of the table. The values above 95% (high enrichment of motifs with respect to proteome) are highlighted in green. The last row lists the number of positive (RK) and negative (ED) residues in the CENP-C protein of a given species.

Cellular component

	Homo sapiens (REF)		upload_1 (Hierarchy) NEW!			
GO cellular component complete	#	#	expected	Fold Enrichment	+/-	P value
pericentric heterochromatin	18	4	.14	28.42	+	1.84E-02
nuclear transcriptional repressor complex	32	6	.25	23.98	+	3.30E-04
SWI/SNF superfamily-type complex	76	9	.59	15.14	+	1.60E-05
heterochromatin	72	6	.56	10.66	+	3.31E-02
histone methyltransferase complex	72	6	.56	10.66	+	3.31E-02
transcriptional repressor complex	87	7	.68	10.29	+	8.68E-03
nuclear speck	198	13	1.55	8.40	+	1.00E-05
nuclear body	355	17	2.78	6.12	+	5.07E-06
nucleoplasm part	739	30	5.78	5.19	+	2.02E-10
chromosome	905	31	7.08	4.38	+	5.91E-09
chromatin	442	15	3.46	4.34	+	3.23E-03
nucleolus	879	22	6.87	3.20	+	2.17E-03
chromosomal part	793	19	6.20	3.06	+	2.14E-02
nucleoplasm	2943	67	23.01	2.91	+	4.90E-14
nuclear lumen	3532	77	27.62	2.79	+	5.25E-16
nuclear part	3929	83	30.72	2.70	+	5.16E-17
intracellular organelle lumen	4397	78	34.38	2.27	+	4.27E-11
organelle lumen	4398	78	34.39	2.27	+	4.32E-11
membrane-enclosed lumen	4398	78	34.39	2.27	+	4.32E-11
intracellular non-membrane-bounded organelle	3890	67	30.42	2.20	+	3.88E-08
non-membrane-bounded organelle	3890	67	30.42	2.20	+	3.88E-08
nucleus	7006	106	54.79	1.93	+	4.20E-13
macromolecular complex	4908	66	38.38	1.72	+	1.57E-03
intracellular organelle part	8198	98	64.11	1.53	+	9.35E-05
organelle part	8402	99	65.70	1.51	+	1.62E-04
intracellular membrane-bounded organelle	10464	113	81.83	1.38	+	8.25E-04
membrane-bounded organelle	12169	128	95.16	1.35	+	7.11E-05
intracellular organelle	12048	125	94.21	1.33	+	5.17E-04
organelle	13117	134	102.57	1.31	+	1.00E-04
intracellular part	13838	133	108.21	1.23	+	1.82E-02
Unclassified	2678	8	20.94	.38	-	0.00E00

Molecular function

	Homo sapiens (REF)	upload_1 (Hierarchy) NEW!				
GO molecular function complete	#	#	expected	Fold Enrichment	+/-	P value
chromo shadow domain binding	6	3	.05	63.94	+	4.53E-02
histone methyltransferase activity (H3-K36 specific)	6	3	.05	63.94	+	4.53E-02
histone methyltransferase activity (H3-K4 specific)	18	6	.14	42.63	+	2.43E-05
lysine-acetylated histone binding	18	5	.14	35.52	+	1.07E-03
histone-lysine N-methyltransferase activity	45	9	.35	25.58	+	3.76E-07
protein-lysine N-methyltransferase activity	57	9	.45	20.19	+	2.91E-06
lysine N-methyltransferase activity	58	9	.45	19.84	+	3.38E-06
histone methyltransferase activity	59	9	.46	19.51	+	3.92E-06
DNA helicase activity	57	8	.45	17.95	+	6.19E-05
protein methyltransferase activity	84	9	.66	13.70	+	7.98E-05
helicase activity	159	17	1.24	13.67	+	4.68E-11
DNA-dependent ATPase activity	86	9	.67	13.38	+	9.73E-05
N-methyltransferase activity	91	9	.71	12.65	+	1.56E-04
purine NTP-dependent helicase activity	105	8	.82	9.74	+	5.99E-03
ATP-dependent helicase activity	105	8	.82	9.74	+	5.99E-03
S-adenosylmethionine-dependent methyltransferase activity	151	11	1.18	9.32	+	1.12E-04
histone binding	160	10	1.25	7.99	+	1.87E-03
methyltransferase activity	222	11	1.74	6.34	+	4.84E-03
transferase activity, transferring one-carbon groups	233	11	1.82	6.04	+	7.65E-03
chromatin binding	499	21	3.90	5.38	+	1.41E-06
ATPase activity, coupled	336	14	2.63	5.33	+	1.42E-03
ATPase activity	452	17	3.53	4.81	+	3.43E-04
nucleoside-triphosphatase activity	792	27	6.19	4.36	+	4.03E-07
pyrophosphatase activity	837	27	6.55	4.13	+	1.34E-06
hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	839	27	6.56	4.12	+	1.41E-06
hydrolase activity, acting on acid anhydrides	842	27	6.58	4.10	+	1.52E-06
poly(A) RNA binding	1171	35	9.16	3.82	+	1.88E-08
zinc ion binding	1202	28	9.40	2.98	+	6.33E-04
RNA binding	1633	38	12.77	2.98	+	2.74E-06
macromolecular complex binding	1355	29	10.60	2.74	+	2.11E-03
DNA binding	2533	53	19.81	2.68	+	2.32E-08
ATP binding	1508	30	11.79	2.54	+	5.90E-03
adenyl ribonucleotide binding	1545	30	12.08	2.48	+	9.58E-03
adenyl nucleotide binding	1556	30	12.17	2.47	+	1.10E-02
transition metal ion binding	1459	28	11.41	2.45	+	2.66E-02
nucleic acid binding	4079	78	31.90	2.45	+	1.29E-12
purine ribonucleoside triphosphate binding	1856	33	14.51	2.27	+	1.86E-02
purine ribonucleoside binding	1866	33	14.59	2.26	+	2.08E-02
purine nucleoside binding	1869	33	14.62	2.26	+	2.15E-02

Table Supplement 2 (Related to Figure 7 and Supplemental Fig. S9). Gene ontology enrichment analysis of RK-clusters.

Proteins in human having at least as many RK-clusters as CENP-C are grouped by molecular functions, and top statistically significant (P<0.05) protein groups by enrichment value are shown.

Fluorophore	**Light source & channel**	**Filter cube**		
		Excitation	**Beamsplitter**	**Emission**
tdEos (red emission)	Colibri / LED555[(1)]	BP550/25[(1)]	FT570[(1)]	BP605/70[(1)]
tdEos (photoconversion)	Colibri / LED405[(1)]	FF01-405/10[(2)]	59004BS[(3)]	59004M[(3)]

Source: [(1)]Zeiss, [(2)]Semrock, [(3)]Chroma

Supplemental Table S3 (Related to Fig. 1). Light sources and filters used for wide field fluorescence imaging.

Supplemental References

Allshire RC, Karpen GH. 2008. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat Rev Genet* **9**: 923-937.

Aravind L, Landsman D. 1998. AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic acids research* **26**: 4413-4421.

Baker RE, Rogers K. 2005. Genetic and genomic analysis of the AT-rich centromere DNA element II of Saccharomyces cerevisiae. *Genetics* **171**: 1463-1475.

Baker RE, Rogers K. 2006. Phylogenetic analysis of fungal centromere H3 proteins. *Genetics* **174**: 1481-1492.

Bao Y, White CL, Luger K. 2006. Nucleosome core particles containing a poly(dA.dT) sequence element exhibit a locally distorted DNA structure. *J Mol Biol* **361**: 617-624.

Beitz E. 2000. TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon. *Bioinformatics* **16**: 135-139.

Berry R, Burnell J. 2005. *The handbook of astronomical image processing*. Willmann-Bell, Richmond, VA.

Biggins S. 2013. The composition, functions, and regulation of the budding yeast kinetochore. *Genetics* **194**: 817-846.

Bishop EP, Rohs R, Parker SC, West SM, Liu P, Mann RS, Honig B, Tullius TD. 2011. A map of minor groove shape and electrostatic potential from hydroxyl radical cleavage patterns of DNA. *ACS chemical biology* **6**: 1314-1320.

Black BE, Cleveland DW. 2011. Epigenetic centromere propagation and the nature of CENP-a nucleosomes. *Cell* **144**: 471-479.

Black BE, Foltz DR, Chakravarthy S, Luger K, Woods VL, Jr., Cleveland DW. 2004. Structural determinants for generating centromeric chromatin. *Nature* **430**: 578-582.

Brown MT. 1995. Sequence similarities between the yeast chromosome segregation protein Mif2 and the mammalian centromere protein CENP-C. *Gene* **160**: 111-116.

Carroll CW, Milks KJ, Straight AF. 2010. Dual recognition of CENP-A nucleosomes is required for centromere assembly. *The Journal of cell biology* **189**: 1143-1155.

Chen Y, Baker RE, Keith KC, Harris K, Stoler S, Fitzgerald-Hayes M. 2000. The N terminus of the centromere H3-like protein Cse4p performs an essential function distinct from that of the histone fold domain. *Mol Cell Biol* **20**: 7037-7048.

Churchill ME, Suzuki M. 1989. 'SPKK' motifs prefer to bind to DNA at A/T-rich sites. *The EMBO journal* **8**: 4189-4195.

Cohen RL, Espelin CW, De Wulf P, Sorger PK, Harrison SC, Simons KT. 2008. Structural and functional dissection of Mif2p, a conserved DNA-binding kinetochore protein. *Molecular biology of the cell* **19**: 4480-4491.

Cole HA, Howard BH, Clark DJ. 2011. The centromeric nucleosome of budding yeast is perfectly positioned and covers the entire centromere. *Proc Natl Acad Sci U S A* **108**: 12687-12692.

Cole JL, Lary JW, T PM, Laue TM. 2008. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol* **84**: 143-179.

Coll M, Frederick CA, Wang AH, Rich A. 1987. A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer

d(CGCAAATTTGCG) and its complex with distamycin. *Proc Natl Acad Sci U S A* **84**: 8385-8389.

Collins KA, Castillo AR, Tatsutani SY, Biggins S. 2005. De novo kinetochore assembly requires the centromeric histone H3 variant. *Molecular biology of the cell* **16**: 5649-5660.

Cooper JL, Henikoff S. 2004. Adaptive evolution of the histone fold domain in centromeric histones. *Mol Biol Evol* **21**: 1712-1718.

Dyer PN, Edayathumangalam RS, White CL, Bao Y, Chakravarthy S, Muthurajan UM, Luger K. 2004. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol* **375**: 23-44.

Earnshaw WC, Rothfield N. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* **91**: 313-321.

Falk SJ, Guo LY, Sekulic N, Smoak EM, Mani T, Logsdon GA, Gupta K, Jansen LE, Van Duyne GD, Vinogradov SA, Lampson MA, Black BE. 2015. Chromosomes. CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. *Science* **348**: 699-703.

Falk SJ, Lee J, Sekulic N, Sennett MA, Lee TH, Black BE. 2016. CENP-C directs a structural transition of CENP-A nucleosomes mainly through sliding of DNA gyres. *Nat Struct Mol Biol*.

Fried M, Crothers DM. 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Research* **9**: 6505-6525.

Fukagawa T. 2004. Centromere DNA, proteins and kinetochore assembly in vertebrate cells. *Chromosome Res* **12**: 557-567.

Fukagawa T, Earnshaw WC. 2014. The centromere: chromatin foundation for the kinetochore machinery. *Dev Cell* **30**: 496-508.

Gascoigne KE, Takeuchi K, Suzuki A, Hori T, Fukagawa T, Cheeseman IM. 2011. Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. *Cell* **145**: 410-422.

Guse A, Carroll CW, Moree B, Fuller CJ, Straight AF. 2011. In vitro centromere and kinetochore assembly on defined chromatin templates. *Nature* **477**: 354-358.

Hegemann JH, Fleig UN. 1993. The centromere of budding yeast. *Bioessays* **15**: 451-460.

Henikoff S, Ahmad K, Malik HS. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* **293**: 1098-1102.

Hong J, Feng H, Zhou Z, Ghirlando R, Bai Y. 2013. Identification of functionally conserved regions in the structure of the chaperone/CenH3/H4 complex. *J Mol Biol* **425**: 536-545.

Hornung P, Troc P, Malvezzi F, Maier M, Demianova Z, Zimniak T, Litos G, Lampert F, Schleiffer A, Brunner M, Mechtler K, Herzog F, Marlovits TC, Westermann S. 2014. A cooperative mechanism drives budding yeast kinetochore assembly downstream of CENP-A. *The Journal of cell biology* **206**: 509-524.

Huerta-Cepas J, Serra F, Bork P. 2016. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Molecular biology and evolution* **33**: 1635-1638.

Huth JR, Bewley CA, Nissen MS, Evans JN, Reeves R, Gronenborn AM, Clore GM. 1997. The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nature structural biology* **4**: 657-665.

Huynh VA, Robinson PJ, Rhodes D. 2005. A method for the in vitro reconstitution of a defined "30 nm" chromatin fibre containing stoichiometric amounts of the linker histone. *J Mol Biol* **345**: 957-968.

Kato H, Jiang J, Zhou BR, Rozendaal M, Feng H, Ghirlando R, Xiao TS, Straight AF, Bai Y. 2013. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* **340**: 1110-1113.

Keith KC, Fitzgerald-Hayes M. 2000. CSE4 genetically interacts with the Saccharomyces cerevisiae centromere DNA elements CDE I and CDE II but not CDE III. Implications for the path of the centromere dna around a cse4p variant nucleosome. *Genetics* **156**: 973-981.

Kong X, Liu J, Li L, Yue L, Zhang L, Jiang H, Xie X, Luo C. 2015. Functional interplay between the RK motif and linker segment dictates Oct4-DNA recognition. *Nucleic Acids Res* **43**: 4381-4392.

Lanini L, McKeon F. 1995. Domains required for CENP-C assembly at the kinetochore. *Molecular biology of the cell* **6**: 1049-1059.

Lawrimore J, Bloom KS, Salmon ED. 2011. Point centromeres contain more than a single centromere-specific Cse4 (CENP-A) nucleosome. *The Journal of cell biology* **195**: 573-582.

Liu-Johnson HN, Gartenberg MR, Crothers DM. 1986. The DNA binding domain and bending angle of E. coli CAP protein. *Cell* **47**: 995-1005.

Logsdon GA, Barrey EJ, Bassett EA, DeNizio JE, Guo LY, Panchenko T, Dawicki-McKenna JM, Heun P, Black BE. 2015. Both tails and the centromere targeting domain of CENP-A are required for centromere establishment. *The Journal of cell biology*.

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**: 251-260.

Meeks-Wagner D, Wood JS, Garvik B, Hartwell LH. 1986. Isolation of two genes that affect mitotic chromosome transmission in S. cerevisiae. *Cell* **44**: 53-63.

Meluh PB, Koshland D. 1995. Evidence that the MIF2 gene of Saccharomyces cerevisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Molecular biology of the cell* **6**: 793-807.

Milks KJ, Moree B, Straight AF. 2009. Dissection of CENP-C-directed centromere and kinetochore assembly. *Molecular biology of the cell* **20**: 4246-4255.

Mizuguchi G, Xiao H, Wisniewski J, Smith MM, Wu C. 2007. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. *Cell* **129**: 1153-1164.

Moroi Y, Peebles C, Fritzler MJ, Steigerwald J, Tan EM. 1980. Autoantibody to centromere (kinetochore) in scleroderma sera. *Proc Natl Acad Sci U S A* **77**: 1627-1631.

Nelson HC, Finch JT, Luisi BF, Klug A. 1987. The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature* **330**: 221-226.

Politi V, Perini G, Trazzi S, Pliss A, Raska I, Earnshaw WC, Della Valle G. 2002. CENP-C binds the alpha-satellite DNA in vivo at specific centromere domains. *J Cell Sci* **115**: 2317-2327.

Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, Farrell CM, Hart J, Landrum MJ, McGarvey KM, Murphy MR, O'Leary NA, Pujar S, Rajput B, Rangwala SH, Riddick LD, Shkeda A, Sun H, Tamez P, Tully RE, Wallin C, Webb D, Weber J, Wu W, DiCuccio M, Kitts P, Maglott DR, Murphy TD, Ostell JM.. 2014. RefSeq: an update on mammalian reference sequences. *Nucleic acids research* **42**: D756-763.

Reeves R. 2000. Structure and function of the HMGI(Y) family of architectural transcription factors. *Environ Health Perspect* **108 Suppl 5**: 803-809.

Reeves R, Nissen MS. 1990. The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J Biol Chem* **265**: 8573-8582.

Rohs R, Jin X, West SM, Joshi R, Honig B, Mann RS. 2010. Origins of specificity in protein-DNA recognition. *Annual review of biochemistry* **79**: 233-269.

Rohs R, West SM, Sosinsky A, Liu P, Mann RS, Honig B. 2009. The role of DNA shape in protein-DNA recognition. *Nature* **461**: 1248-1253.

Saitoh H, Tomkiel J, Cooke CA, Ratrie H, 3rd, Maurer M, Rothfield NF, Earnshaw WC. 1992. CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. *Cell* **70**: 115-125.

Samel A, Cuomo A, Bonaldi T, Ehrenhofer-Murray AE. 2012. Methylation of CenH3 arginine 37 regulates kinetochore integrity and chromosome segregation. *Proc Natl Acad Sci U S A* **109**: 9029-9034.

Schuck P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**: 1606-1619.

Schuck P. 2003. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem* **320**: 104-124.

Sekulic N, Bassett EA, Rogers DJ, Black BE. 2010. The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. *Nature* **467**: 347-351.

Shaytan AK, Armeev GA, Goncearenco A, Zhurkin VB, Landsman D, Panchenko AR. 2016. Coupling between Histone Conformations and DNA Geometry in Nucleosomes on a Microsecond Timescale: Atomistic Insights into Nucleosome Functions. *J Mol Biol* **428**: 221-237.

Shen A, Higgins DE, Panne D. 2009. Recognition of AT-rich DNA binding sites by the MogR repressor. *Structure (London, England : 1993)* **17**: 769-777.

Stefl R, Wu H, Ravindranathan S, Sklenar V, Feigon J. 2004. DNA A-tract bending in three dimensions: solving the dA4T4 vs. dT4A4 conundrum. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 1177-1182.

Sugimoto K, Kuriyama K, Shibata A, Himeno M. 1997. Characterization of internal DNA-binding and C-terminal dimerization domains of human centromere/kinetochore autoantigen CENP-C in vitro: role of DNA-binding and self-associating activities in kinetochore organization. *Chromosome Res* **5**: 132-141.

Sullivan BA, Blower MD, Karpen GH. 2001. Determining centromere identity: cyclical stories and forking paths. *Nat Rev Genet* **2**: 584-596.

Tachiwana H, Kagawa W, Shiga T, Osakabe A, Miya Y, Saito K, Hayashi-Takanaka Y, Oda T, Sato M, Park SY, Kimura H, Kurumizaka H. 2011. Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* **476**: 232-235.

Tachiwana H, Kurumizaka H. 2011. Structure of the CENP-A nucleosome and its implications for centromeric chromatin architecture. *Genes Genet Syst* **86**: 357-364.

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A. 2003. PANTHER: a library of protein families and subfamilies indexed by function. *Genome research* **13**: 2129-2141.

Trazzi S, Perini G, Bernardoni R, Zoli M, Reese JC, Musacchio A, Della Valle G. 2009. The C-terminal domain of CENP-C displays multiple and critical functions for mammalian centromere formation. *PLoS One* **4**: e5832.

Weir JR, Faesen AC, Klare K, Petrovic A, Basilico F, Fischbock J, Pentakota S, Keller J, Pesenti ME, Pan D, Vogt D, Wohlgemuth S, Herzog F, Musacchio A. 2016. Insights from biochemical reconstitution into the architecture of human kinetochores. *Nature* **537**: 249-253.

Westermann S, Drubin DG, Barnes G. 2007. Structures and functions of yeast kinetochore complexes. *Annual review of biochemistry* **76**: 563-591.

Westermann S, Schleiffer A. 2013. Family matters: structural and functional conservation of centromere-associated proteins from yeast to humans. *Trends Cell Biol* **23**: 260-269.

Wisniewski J, Hajj B, Chen J, Mizuguchi G, Xiao H, Wei D, Dahan M, Wu C. 2014. Imaging the fate of histone Cse4 reveals de novo replacement in S phase and subsequent stable residence at centromeres. *Elife* **3**: e02203.

Xiao H, Mizuguchi G, Wisniewski J, Huang Y, Wei D, Wu C. 2011. Nonhistone Scm3 binds to AT-rich DNA to organize atypical centromeric nucleosome of budding yeast. *Mol Cell* **43**: 369-380.

Yang CH, Tomkiel J, Saitoh H, Johnson DH, Earnshaw WC. 1996. Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. *Mol Cell Biol* **16**: 3576-3586.

Zhao H, Brautigam CA, Ghirlando R, Schuck P. 2013a. Overview of current methods in sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. *Curr Protoc Protein Sci* **Chapter 20**: Unit20 12.

Zhao H, Ghirlando R, Piszczek G, Curth U, Brautigam CA, Schuck P. 2013b. Recorded scan times can limit the accuracy of sedimentation coefficients in analytical ultracentrifugation. *Anal Biochem* **437**: 104-108.

Zhou Z, Feng H, Zhou BR, Ghirlando R, Hu K, Zwolak A, Miller Jenkins LM, Xiao H, Tjandra N, Wu C, Bai Y. 2011. Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. *Nature* **472**: 234-237.

Zimarino V, Wu C. 1987. Induction of sequence-specific binding of Drosophila heat shock activator protein without protein synthesis. *Nature* **327**: 727-730.