2 September, 2015

SUPPLEMENTARY DATA

Novel Nucleosomal Particles Containing Core Histones and Linker DNA but no Histone H1

Hope A. Cole, Feng Cui, Josefina Ocampo, Tara L. Burke, Tatiana Nikitina, V. Nagarajavel, Naoe Kotomura, Victor B. Zhurkin and David J. Clark

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Proto-chromatosomes have a 7 or 8-bp extension on one side (154-bp particles) or both sides (161-bp particles) of the nucleosome core in wild type cells. Analysis of less digested nucleosomes (data in Figure 1D). **(A)** Calculation of cross-correlation between the right-hand end of a core particle sequence (red line) and the right-hand end of a proto-chromatosome sequence (blue line). The right-hand end refers to the location of the sequence with respect to the chromosome after alignment. **(B)** Cross-correlation between nucleosome core particles (145-148 bp) and proto-chromatosomes (153-155 bp). **(C)** Cross-correlation between core particles (145-148 bp) and proto-chromatosomes (160-162 bp).

Figure S2. Nucleosome phasing relative to the transcription start site (TSS) in wild type cells after MNase-only or MNase-ExoIII digestion. The midpoints of all nucleosome sequences between 140 and 160 bp long were mapped relative to the TSS, summed for ~4150 genes with a known TSS and then normalized to the average (Cole *et al*., 2014). **(A)** Comparison of MNase only (see Figure 1C) and MNase-ExoIII data (see Figure 1E), using nucleosome sequences between 140 and 160 bp. **(B)** Comparison of proto-chromatosomes (152-156 bp; black line) with core particles (144-149 bp; red line) in MNase-ExoIII data (see Figure 1D).

Figure S3. The strand-specific preference for A *versus* T might involve AA:TT wedges. A view of a segment of nucleosomal DNA at SHL -4.5 from a crystal structure (Davey *et al*., 2002). The green ribbon indicates the A-strand (5'-CCAAAA); the yellow ribbon indicates the T-strand (5'-TTTTGG); the DNA is kinked at the CA:TG step. The T-strand is slightly more extended than the A-strand, resulting in a wedge and facilitating bending towards the histone octamer (blue arrow). The first A is located 27 bp from the border (coinciding with an enhanced A-peak in Figure 5B). Also shown is an arginine residue (H2B-R30), which penetrates the minor groove and is closer to the T-strand (Wang *et al*., 2010).

Figure S4. Distance auto-correlation analysis for MNase-ExoIII nucleosome core particles (145-147 bp) (data in Figure 1E). **(A)** Auto-correlation analysis. **(B)** Regression plot of peak values to determine the average distance between peaks in A. The average spacing is 161 bp, similar to reported repeat lengths for yeast chromatin (discussed by Cui *et al*., 2012). **(C)** Fine detail of the plot in A. *D*. Regression plot of peak values to determine the average distance between peaks in C. The period is 10.18 bp and extends into the first linker. The value is close to the helical repeat of DNA and indicates that the overlapping translational positions adopted by nucleosomes tend to be rotationally related.

Figure S5. A steric clash between ExoIII and the nucleosome core may occur at a location 7 or 8 bp external to the nucleosome core (Davey *et al*., 2002). The DNA has been extended out from

the nucleosome core, assuming a straight trajectory; the 3'-end is indicated by the arrow on the yellow ribbon. The circle indicates the approximate size and location of ExoIII, based on the modelling of DNA binding to ExoIII using the crystal structure of ExoIII (Mol *et al*., 2010), and the fact that the enzyme degrades the 3'-strand (the 5'-strand is left intact and is removed by MNase). At -7/-8, ExoIII is located on the inward facing surface of the linker DNA and its rotation may be inhibited; at -1, where the core particle begins, ExoIII would be bound facing outwards.

SUPPLEMENTARY REFERENCES

Cole,H.A., Ocampo,J., Iben,J.R., Chereji,R.V. and Clark,D.J. (2014) Heavy transcription of yeast genes correlates with differential loss of histone H2B relative to H4 and queued RNA polymerases. *Nucl. Acids Res.,* **42,** 12512-12522.

Cui,F., Cole,H.A., Clark,D.J. and Zhurkin,V.B. (2012) Transcriptional activation of yeast genes disrupts intragenic nucleosome phasing. *Nucl. Acids Res.,* **40,** 10753-10764.

Davey,C.A., Sargent,D.F., Luger,K., Maeder,A.W. and Richmond,T.J. (2002) Solvent-mediated interactions in the structure of the nucleosome core particle at 1.9Å resolution. *J. Mol. Biol.,* **319,** 1097-1113.

Mol,C.D., Kuo,C., Thayer,M.M., Cunningham,R.P. and Tainer,J.A. (1995) Structure and function of the multifunctional DNA-repair enzyme exonuclease III. *Nature,* **374,** 381-386.

Wang,D., Ulyanov,N.B. and Zhurkin,V.B. (2010) Sequence-dependent kink and slide deformations of nucleosomal DNA facilitated by histone arginines bound in the minor groove. *J. Biomol. Struc. Dyn.,* **27,** 843-859.

SUPPLEMENTARY TABLE S1. Summary of paired-end data.

Sample[1]	Digest	Read length	No. of aligned reads	Pearson Correlation, R^3
JRY4012 (WT) -2	MNase only	45[2]	28,604,991	0.89
JRY4012 (WT) -3	MNase only	50	13,027,140	
JRY4012 (WT) -1	MNase-ExoIII (less digested)	50	4,244,868	0.79
JRY4012 (WT) -2	MNase-ExoIII (less digested)	50	11,657,436	
JRY4012 (WT) -1	MNase-ExoIII (limit digest)	50	3,897,676	0.71
JRY4012 (WT) -2	MNase-ExoIII (limit digest)	50	18,768,045	
YVN381 (*hho1Δ*) -1	MNase-ExoIII (less digested)	50	2,306,857	0.64
YVN381 (*hho1Δ*) -2	MNase-ExoIII (less digested)	50	19,962,938	
YVN381 (*hho1Δ*) -1	MNase-ExoIII (limit digest)	50	4,233,939	0.79
YVN381 (*hho1Δ*) -2	MNase-ExoIII (limit digest)	50	8,286,479	
pRS-ARG1 Yeast Recon -1	MNase-ExoIII (less digested)	50	1,718,936	0.91
pRS-ARG1 Yeast Recon -2	MNase-ExoIII (less digested)	50	1,398,403	
pRS-ARG1 Yeast Recon -1	MNase-ExoIII (limit digest)	50	1,381,761	0.85
pRS-ARG1 Yeast Recon -2	MNase-ExoIII (limit digest)	50	1,345,274	
pRS-ARG1 Chick Recon -1	MNase-ExoIII (less digested)	50	3,331,919	0.98
pRS-ARG1 Chick Recon -2	MNase-ExoIII (less digested)	50	4,096,158	
pRS-ARG1 Chick Recon -1	MNase-ExoIII (limit digest)	50	2,586,495	0.99
pRS-ARG1 Chick Recon -2	MNase-ExoIII (limit digest)	50	3,417,780	
Native mouse (N3) -1	MNase-ExoIII	50	44,234,494	
Native mouse (N4) -1	MNase-ExoIII	50	42,344,272	
Native mouse (N5) -1	MNase-ExoIII	50	54,193,618	
H1-depleted mouse (D2) -1	MNase-ExoIII	50	39,405,698	
H1-depleted mouse (D3) -1	MNase-ExoIII	50	54,630,830	
H1-depleted mouse (D4) -1	MNase-ExoIII	50	45,089,544	
Native mouse (N4) -2	MNase-ExoIII	50	44,474,400	
Native mouse (N5) -2	MNase-ExoIII	50	48,228,514	
Native mouse (N6) -2	MNase-ExoIII	50	53,092,154	
H1-depleted mouse (D3) -2	MNase-ExoIII	50	44,842,412	
H1-depleted mouse (D4) -2	MNase-ExoIII	50	54,514,652	
H1-depleted mouse (D5) -2	MNase-ExoIII	50	53,923,926	

[1] Biological replicate experiments are indicated by "-1", "-2" and "-3". WT = wild type. "Recon.": nucleosomes reconstituted *in vitro* using yeast recombinant histones or native chicken erythrocyte histones on plasmid pRS-ARG1-B.
The mouse liver chromatin samples represent a series of digestion points (N or D) with increasing amounts of MNase and fixed ExoIII (see Figure 4).
[2] Reads were 50 nt but included a 5-nt bar code (GTATT).
[3] The Pearson correlation, *R*, was calculated for the nucleosome occupancy profiles as described (Cui *et al*., 2012). Note that *R* is affected by the number of aligned reads if the coverage is relatively low. *R* was not calculated for the mouse samples because the genome coverage is too low to obtain useful occupancy data.

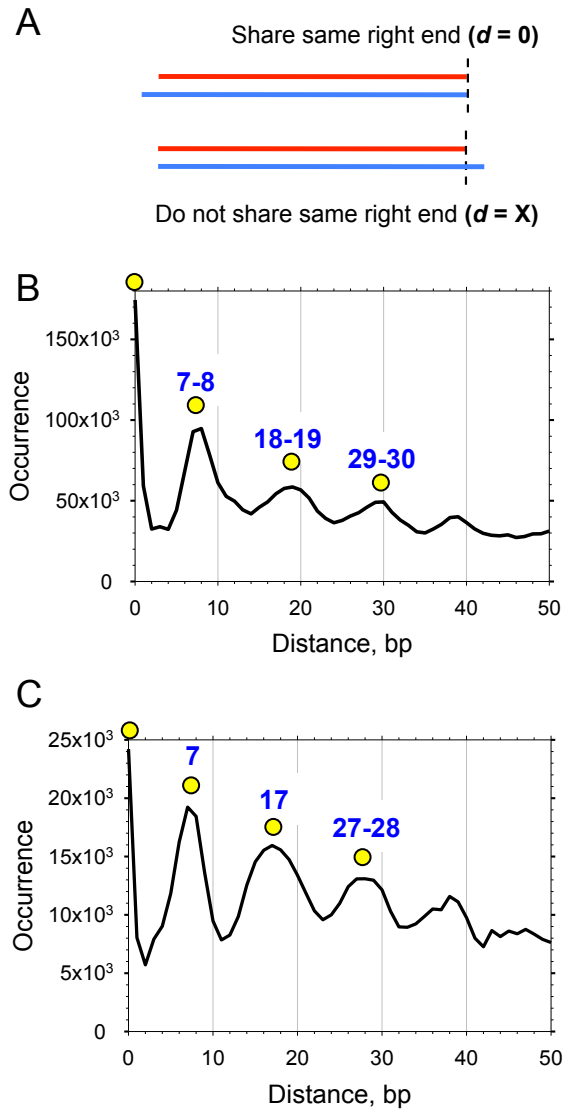

Figure S1. Proto-chromatosomes have a 7 or 8-bp extension on one side (154-bp particles) or both sides (161-bp particles) of the nucleosome core in wild type cells. Analysis of less digested nucleosomes (data in Figure 1D). **(A)** Calculation of cross-correlation between the right-hand end of a core particle sequence (red line) and the right-hand end of a proto-chromatosome sequence (blue line). The right-hand end refers to the location of the sequence with respect to the chromosome after alignment. **(B)** Cross-correlation between nucleosome core particles (145-148 bp) and proto-chromatosomes (153-155 bp). **(C)** Cross-correlation between core particles (145-148 bp) and proto-chromatosomes (160-162 bp).

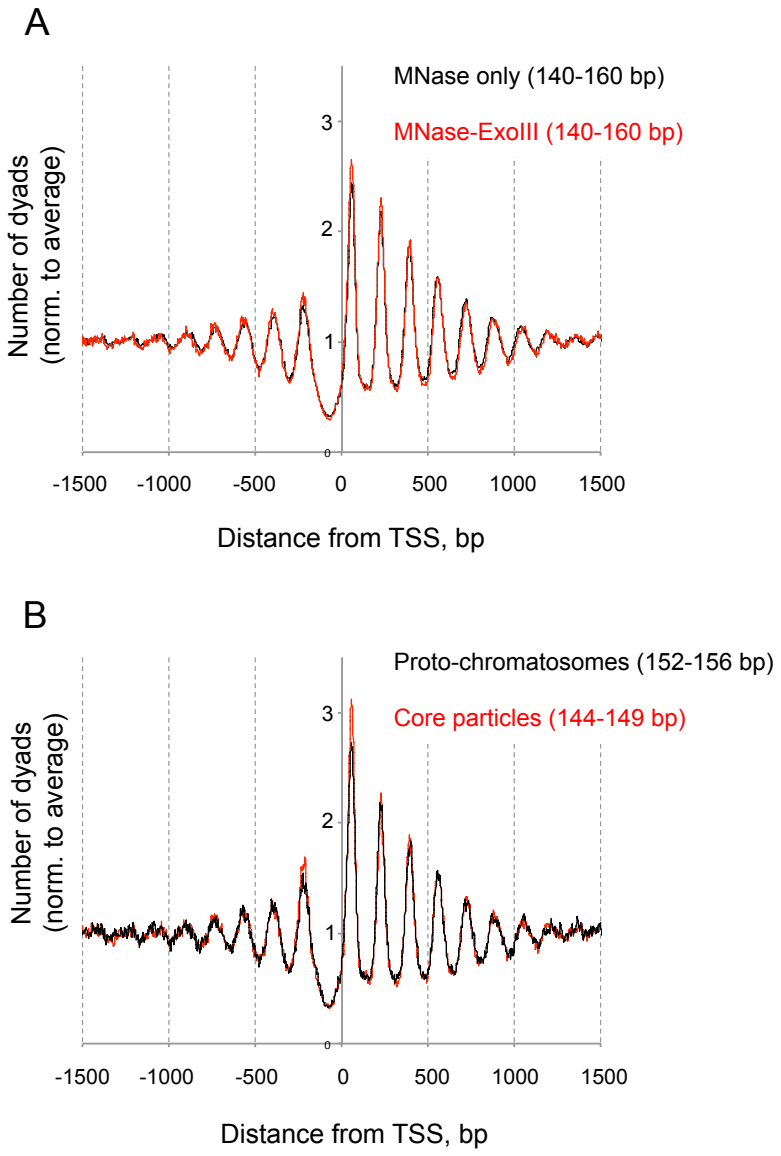

Figure S2. Nucleosome phasing relative to the transcription start site (TSS) in wild type cells after MNase-only or MNase-ExoIII digestion. The midpoints of all nucleosome sequences between 140 and 160 bp long were mapped relative to the TSS, summed for ~4150 genes with a known TSS and then normalized to the average (Cole *et al*., 2014). **(A)** Comparison of MNase only (see Figure 1C) and MNase-ExoIII data (see Figure 1E), using nucleosome sequences between 140 and 160 bp. **(B)** Comparison of proto-chromatosomes (152-156 bp; black line) with core particles (144-149 bp; red line) in MNase-ExoIII data (see Figure 1D).

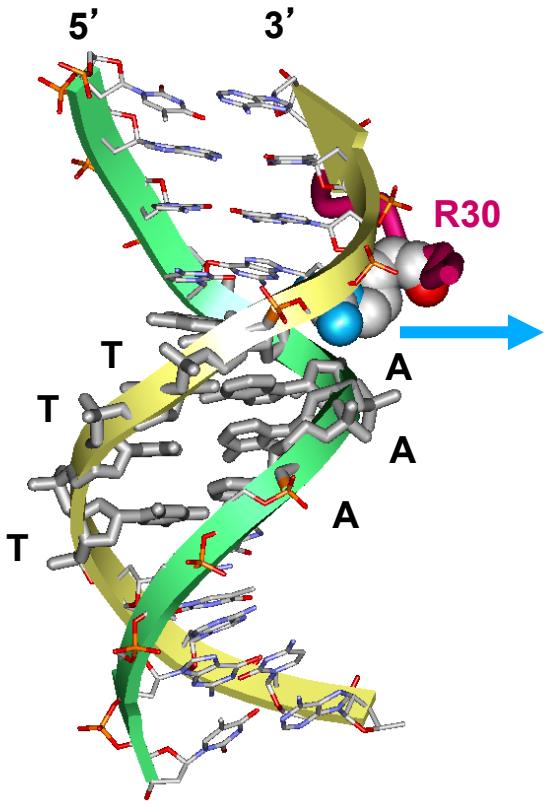

Figure S3. The strand-specific preference for A *versus* T might involve AA:TT wedges. A view of a segment of nucleosomal DNA at SHL -4.5 from a crystal structure (Davey *et al*., 2002). The green ribbon indicates the A-strand (5'-CCAAAA); the yellow ribbon indicates the T-strand (5'-TTTTGG); the DNA is kinked at the CA:TG step. The T-strand is slightly more extended than the A-strand, resulting in a wedge and facilitating bending towards the histone octamer (blue arrow). The first A is located 27 bp from the border (coinciding with an enhanced A-peak in Figure 4B). Also shown is an arginine residue (H2B-R30), which penetrates the minor groove and is closer to the T-strand (Wang *et al*., 2010).

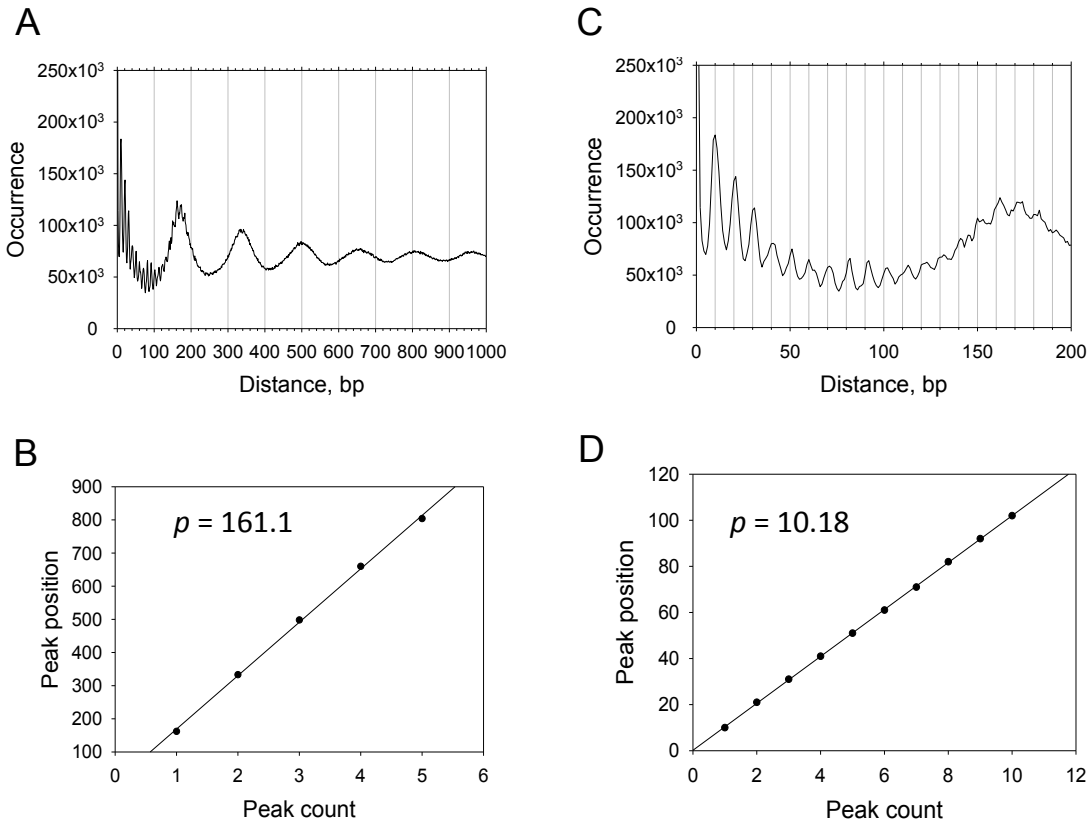

Figure S4. Distance auto-correlation analysis for MNase-ExoIII nucleosome core particles (145-147 bp) (data in Figure 1E). **(A)** Auto-correlation analysis. **(B)** Regression plot of peak values to determine the average distance between peaks in A. The average spacing is 161 bp, similar to reported repeat lengths for yeast chromatin (discussed by Cui *et al*., 2012). **(C)** Fine detail of the plot in A. *D*. Regression plot of peak values to determine the average distance between peaks in C. The period is 10.18 bp and extends into the first linker. The value is close to the helical repeat of DNA and indicates that the overlapping translational positions adopted by nucleosomes tend to be rotationally related.

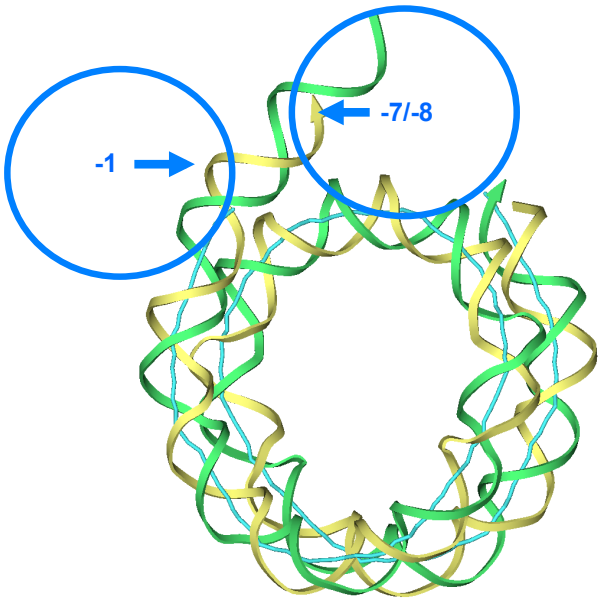

Figure S5. A steric clash between ExoIII and the nucleosome core may occur at a location 7 or 8 bp external to the nucleosome core (Davey *et al*., 2002). The DNA has been extended out from the nucleosome core, assuming a straight trajectory; the 3'-end is indicated by the arrow on the yellow ribbon. The circle indicates the approximate size and location of ExoIII, based on the modelling of DNA binding to ExoIII using the crystal structure of ExoIII (Mol *et al*., 2010), and the fact that the enzyme degrades the 3'-strand (the 5'-strand is left intact and is removed by MNase). At -7/-8, ExoIII is located on the inward facing surface of the linker DNA and its rotation may be inhibited; at -1, where the core particle begins, ExoIII would be bound facing outwards.