7882–7899 *Nucleic Acids Research, 2023, Vol. 51, No. 15* *Published online 10 July 2023*
https://doi.org/10.1093/nar/gkad577

Histone divergence in trypanosomes results in unique alterations to nucleosome structure

Gauri Deák[1], Hannah Wapenaar[1], Gorka Sandoval[1], Ruofan Chen[1], Mark R.D. Taylor[1], Hayden Burdett[1], James A. Watson[1], Maarten W. Tuijtel[1,2], Shaun Webb[1] and Marcus D. Wilson[1,*]

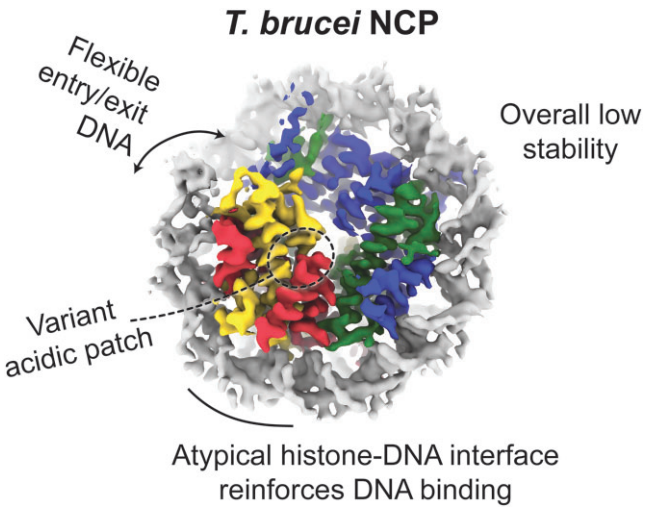
[1]Wellcome Centre for Cell Biology, University of Edinburgh, Michael Swann Building, Kings Buildings, Mayfield Road, Edinburgh EH9 3JR, UK and [2]Department of Molecular Sociology, Max Planck Institute of Biophysics, Max-von-Laue-Straße 3, 60438 Frankfurt am Main, Germany

Received March 13, 2023; Revised June 02, 2023; Editorial Decision June 20, 2023; Accepted June 26, 2023

ABSTRACT

Eukaryotes have a multitude of diverse mechanisms for organising and using their genomes, but the histones that make up chromatin are highly conserved. Unusually, histones from kinetoplastids are highly divergent. The structural and functional consequences of this variation are unknown. Here, we have biochemically and structurally characterised nucleosome core particles (NCPs) from the kinetoplastid parasite *Trypanosoma brucei*. A structure of the *T. brucei* NCP reveals that global histone architecture is conserved, but specific sequence alterations lead to distinct DNA and protein interaction interfaces. The *T. brucei* NCP is unstable and has weakened overall DNA binding. However, dramatic changes at the H2A-H2B interface introduce local reinforcement of DNA contacts. The *T. brucei* acidic patch has altered topology and is refractory to known binders, indicating that the nature of chromatin interactions in *T. brucei* may be unique. Overall, our results provide a detailed molecular basis for understanding evolutionary divergence in chromatin structure.

GRAPHICAL ABSTRACT



INTRODUCTION

Nucleosomes are the basic unit of chromatin and control chromatin-associated processes in eukaryotic genomes. The nucleosome core particle (NCP) is composed of DNA wrapped around an octamer of four histone proteins (H2A, H2B, H3 and H4) and serves as a DNA compaction unit (1,2), inhibitor of transcription (3) and dynamic molecular interaction platform (4). In line with their role as architectural proteins, histone sequences are highly conserved, especially in well-studied eukaryotes (5,6). NCP structures from vertebrates (7,8), invertebrates (9) and unicellular eukaryotes such as yeasts (10,11) have revealed remarkable conservation in global histone architecture and key sites of histone-DNA interactions. However, even small changes to histone primary sequences can have large structural and functional consequences on NCP structure (8,12–14).

Histone sequence variation and nucleosome structure in highly divergent eukaryotes are relatively understudied. The

*To whom correspondence should be addressed. Tel: +44 1316507029; Email: marcus.wilson@ed.ac.uk

group Kinetoplastida is ranked amongst the most evolutionarily ancestral groups of parasitic protists and was estimated to have split from other eukaryotic lineages around 500 million years ago (15). Kinetoplastida includes multiple pathogens, particularly those belonging to the *Trypanosoma* and *Leishmania* species. Of these, *Trypanosoma brucei* is a major clinical target, causing both human and animal trypanosomiasis (16–18). In *T. brucei,* chromatin accessibility has a direct effect on antigenic variation, a key immune evasion mechanism contributing to its pathogenicity (19). Further research on understanding how the trypanosome genome is organised and responds to stimuli could therefore have direct clinical and economic benefits.

Trypanosome chromatin has a number of unusual features. The *T. brucei* genome is organised into 11 large megabase chromosomes, a small, varying number of intermediate-sized chromosomes, and ∼100 minichromosomes (20). Unlike in Metazoa, mitotic chromosome compaction levels in *T. brucei* are low (∼1.2-fold) (21). Most genes are arranged in intron-less polycistronic transcription units (20) and the chromatin context of transcription initiation and termination is in part defined by histones (22–24). *T. brucei* histones have been found to be highly divergent (25), and novel trypanosome-specific histone post-translational modifications (24,26) and chromatin interactors (27–30) have been identified. However, the molecular details of how local nucleosome-level chromatin structure and molecular pathways in the nucleus intersect have been largely unexplored.

Here, we present the cryo-EM structure of the *T. brucei* NCP. The structure reveals altered histone–DNA contact sites, histone-histone interactions, and exposed histone surfaces compared to well-studied model eukaryotes. Globally, *T. brucei* NCPs are unstable and have reduced DNA binding. However, this instability is partly compensated by dramatic alterations in the electrostatic properties of *T. brucei* NCPs at the H2A–H2B binding interface. Furthermore, the surface topology, charge distribution, and binding properties of the *T. brucei* acidic patch are altered. Our phylogenetic analysis of kinetoplastid histones and their predicted structures reveals that these differences identified for *T. brucei* NCPs are likely conserved across the kinetoplastids. Overall, our study provides a molecular basis for understanding and further exploring how DNA compaction and chromatin interactions occur in *T. brucei*.

MATERIALS AND METHODS

Generation of histone plasmid constructs

Plasmids encoding histones from *T. brucei brucei* (protein sequences corresponding to H2A: Tb927.7.2820, H2B: Tb927.10.10480, H3: Tb927.1.2430 and H4: Tb927.5.4170) were received as a gift from the Janzen lab (31). We refer to the species *T. brucei brucei* as '*T. brucei*' only throughout the study. Histone mutations were made either using site directed mutagenesis or cloning of synthesised double-stranded gBlock fragments containing mutations (Integrated DNA Technologies). Tailless *T. brucei* histones (H2A aa. 11–121, H2B aa. 15–111, H3 aa. 24–132, and H4 aa. 18–99) were designed based on sequence alignments to tailless *Xenopus laevis* histones that were previously obtained as products of trypsin digestion (32). *T. brucei* H2B tailless was cloned with an N-terminal His_6-TEV tag to improve protein expression.

Histone protein purification

Histones were expressed in BL-21 DE3 RIL cells and purified from inclusion bodies essentially as described previously (31,33–35) with some modifications. Briefly, for all *Homo sapiens* histones and *T. brucei* H2A, H3 and variants, inclusion bodies were extensively washed (50 mM Tris pH7.5, 100 mM NaCl, 1 mM EDTA, 1 mM benzamidine, 5 mM BME), disrupted in DMSO, and resolubilised (7 M guanidine–HCl, 20 mM Tris pH 7.5, 5 mM DTT). They were then dialysed into urea dialysis buffer (7 M urea, 100 mM NaCl, 15 mM Tris pH 7.5, 1 mM EDTA, 5 mM BME) for 12 h. Dialysis and the cation chromatography step for *T. brucei* histones H2B, H2B tailless, and H4 were performed at reduced pH 7. Purified histones were dialysed into 2 mM BME for 12 h, lyophilized, and stored at –20°C. Human NCPs were assembled using sequences corresponding to human H2A.1, H2B.1, H3.1 C96S C110A and H4 and are referred to as *H. sapiens* NCPs throughout.

The His_6-TEV tag on *T. brucei* H2B tailless was cleaved by overnight incubation of the lyophilized histone with His-TEV protease (1/100 TEV:histone mass ratio) in 1M urea, 100 mM NaCl, 20 mM Tris pH 7.5, 4 mM sodium citrate and 2 mM BME. The cleaved histone was then collected as flow-through from a nickel affinity column (HiTrap IMAC HP, Cytiva) in 1M urea, 100 mM NaCl, 20 mM Tris pH 7.5, 4 mM sodium citrate, 10 mM imidazole, and 2 mM BME.

Histone protein mass was confirmed by 1D intact weight ESI mass spectrometry (SIRCAMs, School of Chemistry, University of Edinburgh) (Supplementary Figure S2A). Concentrations were determined via absorbance at 280 nm using a Nanodrop One spectrophotometer (Thermo Scientific).

GST-PFV-GAG protein purification

A peptide derived from Prototype Foamy Virus (PFV) GAG (UniProt ID: P14349, aa. 535–550) that binds to the acidic patch (36) was cloned into a pET His_6-GST-TEV-LIC plasmid vector and expressed recombinantly in *E. coli* BL21 DE3 RIL cells using overnight induction at 18°C with 0.4 mM IPTG. The protein was then purified using nickel affinity chromatography (HiTrap IMAC HP, Cytiva) (20 mM Tris pH7.5, 400 mM NaCl, 10% glycerol, 2 mM BME, 15–400 mM imidazole gradient) and size exclusion chromatography (HiLoad 16/600 Superdex 200, GE Healthcare) (150 mM NaCl, 5% glycerol, 15 mM HEPES, 2 mM DTT).

NCP reconstitution

NCPs were reconstituted essentially as described (33,34) with some alterations to improve stability. Briefly, DNA for wrapping all NCPs except for hydroxyl radical footprinting assays was generated by isolating large-scale quantities of the plasmids pUC57 8 × 145 bp Widom-601 DNA or 32 × 147 bp alpha-satellite DNA by multiple rounds of

MaxiPrep Kit purifications (Qiagen). The 145 bp fragments were digested and extracted from the plasmid using EcoRV digestion and subsequent PEG and ethanol precipitation steps. For hydroxyl radical footprinting, linker DNA was required to avoid initial undigested signal and the 175bp Widom-601 sequence was used. Fluorescently-tagged DNA was generated by PCR essentially as described (35,37). *T. brucei* 147 bp centromere-associated repeat DNA and 177 bp minichromosome DNA sequences were obtained as gblock fragments (Integrated DNA Technologies), cloned into pUC57 plasmid vectors, amplified by PCR, and purified as described (35,37). All primers and sequences used can be found in Supplementary Table S2.

Purified octamers were wrapped with the DNA using an 18 h exponential salt reduction gradient. The extent and purity of NCP wrapping was checked by native PAGE and SDS-PAGE analysis (Supplementary Figure S2B-D). Due to the appearance of a higher molecular weight species in *T. brucei* NCPs, different molar ratios of octamer:DNA (0.4–1.2, were [DNA] = 0.3 mg/ml) were tested to optimize wrapping efficiency (Supplementary Figure S2B-C). Where necessary, NCPs were purified on HiLoad 16/600 Superdex 200 size exclusion column (GE Healthcare) in 20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT to enrich for NCP only fractions. NCPs were then dialysed for 3 h into a customised Storage Buffer (25 mM NaCl, 2.5% glycerol (v/v), 15 mM HEPES pH 7.5, 1 mM DTT), concentrated, and stored at 4°C for maximum of 1 month. For all biochemical experiments, *H. sapiens* and *T. brucei* NCPs were treated identically and processed concurrently.

Cryo-EM grid preparation and transmission electron microscopy

For cryo-EM grid preparation, NCPs were diluted to a final DNA concentration of 110 μg/ml (DNA concentration) and NaCl concentration of 50 mM. Glutaraldehyde crosslinking agent was added (0.05%) and incubated on ice for 5 min. The reaction was quenched with excess ammonium bicarbonate and Tris pH8. NCPs were concentred through a 100 kDa spin concentrator column (Amicon® Devices) and loaded on an HiLoad 16/600 Superdex 200 size exclusion column (GE Healthcare) in 20 mM HEPEs pH 7.5, 150 mM NaCl, 1mM DTT. Fractions enriched for NCPs were pooled and concentrated (Supplementary Figure S3C).

Monodispersity of the sample was confirmed by negative staining as described (38). Briefly, 5 μg/ml NCPs were applied to 300 mesh copper-grids with continuous carbon-film (C267, TAAB) and stained with 2% uranyl acetate for 2 min prior to washing. Grids were loaded and imaged in F20 TEM operated at 200 kV. Images were collected manually using the EMMENU software (TVIPS) on a TemCam F816 camera (TVIPS) (University of Edinburgh, Transmission EM facility) (Supplementary Figure S3A).

For single particle cryogenic electron microscopy, 3.5 μl of freshly purified and crosslinked NCPs were applied to glow discharged holey carbon quantifoil R 2/2 grids at a concentration of 2.2 μM. Grids were incubated and blotted at 100% humidity and 4°C in a vitrobot mark IV, prior to vitrification in liquid ethane and storage in liquid nitrogen. Grids were screened for ice quality and a small dataset was collected and processed to 2D classes on a TF20 microscope (University of Edinburgh, Cryo-transmission EM facility). Data collection was then performed on a Titan Krios operated at 300 kV equipped with a Gatan K3, operating in correlated double sampling mode. 4193 Lzw compressed tiff movies were obtained using automated serialEM software (39) using a pixel size of 0.829 Å and a total dose of 45.7 electrons/Å^2 (Supplementary Table S1).

Cryo-EM image processing

All micrographs were motion-corrected using MotionCor2, removing the first frame. CTF parameters were estimated using patch CTF in cryoSPARC (40) and poor micrographs were discarded. ∼1000 particles were picked manually and 2D classified to produce templates for template-based picking in cryoSPARC. Two rounds of 2D classification were performed to discard poorly averaged particles and discernible secondary-structure features were pooled. The selected classes were used for ab-initio reconstruction and separated into two ab-initio classes. The best class comprising 306 475 particles were re-extracted with a 316 pixel voxel size and subjected to local CTF refinement and homogeneous refinement with a dynamic mask starting at a resolution of 20 Å and yielding a final map at 3.28 Å resolution. This map was used for all model building and figure preparation. Non-uniform refinement (41) was performed, removing some noise and yielding a GS-FSC map of 3.22 Å. C2 symmetry was applied in homogenous refinement. These maps were used only to aid map interpretability during model building. Map quality and anisotropy were assessed manually in Chimera (42) and using 3D-FSC (43). 3D classification was performed using multiple starting classes using the heterogeneous refinement job in cryoSPARC (Figure 2D).

Model building

The crystal structure of the 145 bp Widom-601 DNA was used from PDB: 3LZ0 (44) in the most logically fitting orientation based on Widom 601 DNA asymmetry and best model to map fits. Initial models for *T. brucei* H2A and H2B histones were generated as a dimeric assembly and *T. brucei* H3 and H4 as a tetrameric assembly with glycine linkers in Alphafold2 (45) and docked in the EM map using UCSF ChimeraX (46). The model was adjusted using Coot (47) and ISOLDE (48). Sequences outside of the density were removed manually in Coot and refined using using Phenix real space refinement (49). Protein geometry was assessed with MolProbity (50). Model fit was assessed using map-to-model cross correlations (48) and EMringer (51). Models have reasonable stereochemistry and are in good agreement with the EM density maps. Figures were prepared in UCSF Chimera (42) and ChimeraX (46).

Small angle X-ray scattering (SAXS)

SEC-SAXS experiments were performed at Diamond Light Source on the B21 beamline. Freshly prepared *H. sapiens* and *T. brucei* NCPs were loaded on column at 2.4 mg/ml

(quantification based on total NCP concentration) and separated by an S200 Increase 3.2 size exclusion column in 20mM HEPEs pH 7.5, 150 mM NaCl, and 1 mM DTT prior to injection into the beamline and recorded with 3 s exposure. Data was reduced and analysed using ScÅtter IV (52).

Protein structure analysis

Structural alignments between *T.* brucei and *H. sapiens* histones were generated using UCSF Chimera or ChimeraX (42,46). Hydrogen bonding interactions between the histones and DNA in the *T. brucei* structure were calculated using PISA (53). The overall molecular dipole moments of *T. brucei* and *H. sapiens* histone octamers were predicted using the Protein Dipole Moments Server (54). The N- and C-terminal tails of *H. sapiens* histones were truncated for this purpose based on the *T. brucei* structure. pK_a values of residues in both *T. brucei* and *H. sapiens* H2A–H2B dimers at SHL3.5 were predicted using PROPKA (55) (Supplementary Data File 1). The surface area of the acidic patch was estimated using PyMOL.

Protein sequence and phylogenetic analysis

To generate phylogenetic trees for each histone, sequences from 20 different organisms sampling kinetoplastids and other eukaryotes were collected from TriTrypDB (56) or NCBI Protein, respectively (57). Multiple sequence alignments were performed with MAFFT (58) and visualized with Jalview (59). Maximum likelihood phylogenetic trees were estimated using IQ-TREE (60), rooted at midpoint, and visualised with iTOL (61). Heatmaps showing percentage identity to *T. brucei* were generated with iTOL using percentage identity matrices calculated by MUSCLE (62). Pairwise percentage identities and similarities between *H. sapiens* and *T. brucei* histones were computed using EMBOSS Needle (63). Isoelectric point (p*I*) values were obtained from Protparam (64).

To obtain larger percentage identity matrices, the TriTrypDB (56) and NCBI Protein (57) databases were mined for histone sequences from 41 organisms (of which 22 were primarily from kinetoplastid reference genomes) and the matrices calculated using MUSCLE. The results were then displayed as correlation maps coloured from 40–100% sequence identity using ggplot2.

Thermal denaturation assays

50 μl reactions with 0.5 μM NCPs, 20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM DTT and 5× SYPRO orange (Life Technologies) were set up in a 96-well plate format and heated from 45°C to 95°C with 0.5°C increments on a Biometra TOptical RT-PCR device (excitation/emission λ = 490/580 nm). Relative fluorescence intensity was normalized as (RFU-RFUmin)/(RFUmax-RFUmin). Tetrasome data was normalized from 60°C due to inherent background signal at lower temperatures. Results from three independent experiments, each with two technical repeats were used to calculate T_m values for *H. sapiens* and *T. brucei* NCPs (Supplementary Data File 2).

Salt stability assays

250 ng of *T. brucei* and *H. sapiens* NCPs (A260 DNA-based quantification) were incubated at various NaCl concentrations (0.5, 1.0, 1.5 and 2M) for 1 h in 10 μl reactions on ice (2.5% glycerol, 15 mM HEPES pH 7.5, 1 mM DTT). After 1 h, NaCl concentrations were normalized to 0.15 M and 22 ng of each sample was loaded onto a 5% Tris-glycine polyacrylamide non-denaturing gel. A control sample kept in Storage Buffer (see 'NCP Reconstitution') was adjacently loaded. Percentage of wrapped NCPs vs. DNA was quantified using the BioRad Image Lab software by comparing the relative contribution of the NCP band and DNA band (summed together as 100%) in each lane. The assay was performed in triplicate for each type of NCP using independently-wrapped NCPs. Quantification can be accessed in Supplementary Data File 3.

Hydroxyl radical footprinting

10 μl samples containing 500 ng of fluorescently labelled nucleosomes (5′ 6-FAM labelled reverse strand, 5′-TAMRA labelled forward strand) were set up in reaction buffer (15 mM HEPES pH 7.5, 25 mM NaCl, 1 mM EDTA, 1 mM DTT). 2.5 μl each of 2 mM ammonium iron (II) sulfate/4 mM EDTA, 0.1 M sodium ascorbate, and 0.12% H_2O_2 were pipetted onto the sides of the reaction tube, mixed together and added to the sample. The reaction was stopped after 4 min by the addition of 100 μl STOP buffer (100 mM Tris pH 7.5, 1% glycerol, 325 mM EDTA, 0.1% SDS, 0.1 mg/ml ProteinaseK [Thermo]). The stopped reaction was then incubated for 20 min at 56°C to allow ProteinaseK digestion to occur. Fragmented DNA was purified by ethanol precipitation and resupended in 10 μl HiDi Formamide. 0.5 μl of GeneScan 500 LIZ size standard (Thermo) was added as a size marker. The resuspended DNA was run on either a 3130xl Genetic or 3730xl DNA Analyzer, operated using the G5 dye filter set. Peaks were analysed using Thermofisher Connect Microsatellite analysis software. Peak size in base pairs were called by the Global southern method.

Micrococcal nuclease (MNase) digestion assays

1.2 μg of NCPs (A260 DNA-based quantification) and 7.2 units of MNase were incubated in 70 μl reactions at 37°C (50 mM Tris pH 8.0, 2.5% glycerol, 25 mM NaCl, 5 mM $CaCl_2$, 1.5 mM DTT). The reaction was stopped by mixing 10 μl at relevant timepoints with 5 μl of stop solution (20 mM Tris pH 8.0, 80 mM EDTA, 80 mM EGTA, 0.25% SDS, 0.5 mg/ml Proteinase K). 44 ng of reaction products were then analysed on a non-denaturing, 5% polyacrylamide TBE gel and stained with Diamond DNA stain (Promega). Experiments were performed with *H. sapiens, T. brucei* WT and *T. brucei Hs*H3 NCPs in triplicate and the disappearance of NCP band intensities were quantified with the BioRad Image Lab software. Quantification can be accessed in Supplementary Data File 4.

Preparing a custom DNA ladder for MNase products

50 μl restriction digest reactions were set up with 2 μg of 145 bp Widom 601 DNA and HpaII, PmlI, or HpaII + PmlI

(1× NEB CutSmart Buffer, 5 U enzyme/μg DNA) for 1 h at 37°C. Restriction enzymes were heat inactivated at 80°C for 2 min. Widom 601 171 bp DNA was generated using PCR as described in the NCP reconstitution section. The final ladder comprised five larger fragments (171, 145, 135, 124 and 114 bp). The fragments were mixed at an equal DNA mass ratio and an aliquot of the final mixture (∼20 ng of each fragment) was loaded on a non-denaturing 5% polyacrylamide gel.

Sequencing MNase reactions

3 μg of NCPs (A260 DNA-based quantification) and 18 U of MNase were incubated in 150 μl reactions for 30 min at 37°C (50 mM Tris pH 8.0, 2.5% glycerol, 25 mM NaCl, 5 mM $CaCl_2$, 1.5 mM DTT). Reactions were quenched with 75 μl of stop solution (20 mM Tris pH 8.0, 80 mM EDTA, 80 mM EGTA, 0.25% SDS, 0.5 mg/ml Proteinase K). Around 1μg of DNA was isolated from each reaction using a Monarch PCR & DNA Cleanup Kit (elution volume of 15 μl, NEB). The DNA was treated with 5 U of Antarctic Phosphatase (NEB #M0289S) for 1 h at 37°C in 20 μl reactions. The phosphatase was heat-inactivated for 2 min at 37°C and the reactions sent for next generation sequencing using the Azenta Amplicon-EZ service (MiSeq 2 × 250).

MNase sequencing data processing

168301 and 132207 reads were obtained for *H. sapiens* and *T. brucei* respectively, with mean Illumina quality scores >30. Adapters were trimmed using Cutadapt (65), aligned to the Widom 601 145 bp sequence using the Burrows-Wheeler Aligner (66), and filtered for length ≤145 bp (0.087% of reads escaped adapter trimming). The final number of mapped reads was 99 212 and 99 050 for *H. sapiens* and *T. brucei* respectively. Read start/end positions extracted from the final dataset are available in Supplementary Data File 4.

Exonuclease III assays

Exonuclease assays were preformed essentially as described (32,67). Briefly, 2 Units of Exonuclease III (Takara) were added to 1 ug of NCPs in ExoIII digestion buffer (50 mM Tris–HCl (pH 8.0), 5 mM $MgCl_2$, 150 mM NaCl and 1 mM DTT). The reaction was incubated at 25°C and samples quenched in stop buffer (20 mM Tris pH 8, 200 mM NaCl, 0.5% SDS, 25 mM EDTA) at regular intervals. DNA products were deproteinized by digestion with 30 ug of proteinase K followed by ethanol precipitation. Samples were processed identically and resuspended in equal volumes of HiDi Formamide, prior to running on a denaturing urea 10% polyacrylamide gel and stained with Diamond DNA stain (Promega). The experiment was repeated in triplicate and quantified using BioRad Image Lab software, quantifying the disappearance of full-length uncut band versus the 0 timepoint taken prior to addition of Exonuclease.

Fluorescence anisotropy peptide binding assays

NCP-binding FITC-labelled peptides derived from Kaposi's Sarcoma Herpesvirus Latency Associated Nuclear Antigen (LANA) (68) and PFV-GAG (36) were synthesised to > 95% purity by BioMatik, Canada. The LANA peptide with mutations L8A R9A S10A ('LANA LRS') and the PFV-GAG peptide with the mutation R540Q ('PFV-GAG RQ') were also synthesized and used as non-binding controls. The peptide sequences used are given below:

LANA (UniProt ID: Q9DUM3, aa. 4–23): FITC-Ahx-PGMRLRSGRSTGAPLTRGSC-Amidation

LANA LRS (UniProt ID: Q9DUM3, aa. 4–23): FITC-Ahx-PGMRAAAGRSTGAPLTRGSC-Amidation

PFV-GAG (UniProt ID: P14349, aa. 535–550) : FITC-Ahx-GGYNLRPRTYQPQRYGG-Amidation

PFV-GAG RQ (UniProt ID: P14349, aa. 535–550): FITC-Ahx-GGYNLQPRTYQPQRYGG-Amidation

Fluorescence anisotropy assays were performed essentially as described (69). 50 nM of peptide tracer was incubated with an increasing concentration of *T. brucei* or *H. sapiens* NCPs from 12.5 nM to 2.4 μM (corresponding to 25 nM to 4.8 μM NCP binding sites) in 20 mM HEPEs pH7.5, 150 mM NaCl, 0.5 mM EDTA, 1 mM DTT and 0.05% Triton X-100. Assays were performed in 25 μl reactions in a 384-well plate and incubated covered at 20°C for 30 min. Fluorescence polarization was measured in an M5 multimode plate reader (Molecular Devices), with 480 nm excitation and 540 nm polarised filters (cut-off at 530 nm). Anisotropy (r) was calculated as below, where the grating factor (G) was approximated to 1, based on FITC alone measurements:

$$r = \frac{I_{VV} - G \times I_{VH}}{I_{VV} + 2G \times I_{VH}}$$

I_{VV} = fluorescence intensity of vertically polarized light

I_{VH} = fluorescence intensity of horizontally polarized light

The resulting values were background subtracted (no NCPs) and plotted against the number of binding sites on the NCPs (2 protein-binding faces per NCP). The experiment was performed in triplicate (Supplementary Data File 5). The resulting graph was fitted using GraphPad Prism with a non-linear regression, saturation binding curve assuming one site with total and non-specific binding. Binding affinity measurement was estimated using the equation:

$$Anisotropy = \frac{B_{max} \times [NCP\ binding\ sites]}{[NCP\ binding\ sites] + K_D} + NS \times [NCP\ binding\ sites]$$

$[NCP\ binding\ sites]$ = molar concentration binding sites in μM (1 on each face of 1 NCP)

B_{max} = maximum anisotropy from specific binding

K_D = equilibrium dissociation constant

NS = slope of non-specific binding (μM^{-1})

Electrophoretic mobility shift assay for acidic patch binding

30 ng of NCPs (DNA-based quantification) were incubated with different concentrations of the purified GST-His-TEV-tagged PFV GAG (0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4 and 1.6 μM) for 30 min in 15 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT, 0.5 mM EDTA, 2.5% glycerol and 0.05 mg/ml BSA at 4°C. Reaction products (20 ng) were then

analysed on a non-denaturing, 5% polyacrylamide TBE gel. Gels were stained with Diamond DNA stain (Promega) and the disappearance of the NCP band quantified using ImageLab. Quantification data can be accessed in Supplementary Data File 5.

RESULTS

Kinetoplastid histones are highly divergent

Previous reports have shown that a subset of trypanosomatid histones are highly divergent compared to model eukaryotes (25). We performed an extended analysis of histone sequences from 22 distinct kinetoplastid genomes including multiple clades, such as those from the *Trypanosoma, Leishmania, Endotrypanum, Crithidia, Angomonas* and *Perkinsela* species (70). This analysis revealed a clear evolutionary divide between the kinetoplastids and a wide sample of eukaryotic taxa (Supplementary Figure S1A). The divide was apparent across all histones, particularly H2A and H2B (Figure 1A). Within the kinetoplastids, our analysis pointed to multiple sub-groups of conserved sequences including segregation between the *Trypanosoma* and *Leishmania spp.* (Supplementary Figure S1A). Aligned to *Homo sapiens* histones, the sequence identities of *T. brucei* histones was low, ranging from 40 to 60% (Figure S1A). Low conservation was apparent even when the predicted unstructured (and more commonly divergent) N- and C-terminal tails were excluded (Figure 1A). This prompted us to investigate whether histone sequence divergence in *T. brucei* leads to functional differences in nucleosome structure, assembly, and function.

Reconstituting the *T. brucei* nucleosome core particle (NCP)

To understand the basic unit of chromatin in *T. brucei,* we used an *in vitro* reconstitution approach to assemble recombinant *T. brucei* nucleosome core particles (NCPs). The four core histones were expressed and purified from *Escherichia coli* (Figure 1B & Supplementary Figure S2A), refolded into octamers, and wrapped with the strong positioning Widom-601 DNA sequence using standard salt dialysis protocols (31,33,34). The process of wrapping octamers with DNA required optimisation due to the presence of a soluble higher molecular weight species (Supplementary Figure S2B–D, see Materials and Methods). We found that the thermal stability of *T. brucei* NCPs was significantly reduced compared to that of *H. sapiens* NCPs (Figure 1C). Wrapping *T. brucei* histone octamers with 147 bp alpha-satellite DNA, another 'strong' positioning sequence (7), also led to more unstable NCPs (Supplementary Figure S2E).

On a coarse level, *H. sapiens* and *T. brucei* NCPs appeared structurally similar. By negative stain transmission electron microscopy, the *T. brucei* NCPs appeared as ∼10 nm disk shapes, reminiscent of NCPs from other species (Supplementary Figure S3A). Furthermore, similar dimensions for *T. brucei* and *H. sapiens* NCPs were obtained by in solution small angle X-ray scattering (SAXS) (Supplementary Figure S3B), suggesting that despite its differences in stability, at low resolutions, *T. brucei* nucleosome structure is maintained.

The cryo-EM structure of *T. Brucei* NCP reveals compressed histone architecture

In order to investigate the molecular details of divergence in *T. brucei* nucleosomes, we determined the structure of *T. brucei* NCP at a global resolution of 3.3 Å by single particle cryogenic electron microscopy (cryo-EM) (Figure 1D, Supplementary Figure S3, Supplementary Movie S1). The low stability of the NCPs required us to perform mild chemical crosslinking prior to purification and sample preparation (Supplementary Figure S3C). The resolution was sufficient to allow us to build a model into the EM density for the core of the NCP (Figure 2A, Supplementary Figure S4A, Supplementary Table S1).

The *T. brucei* NCP forms a characteristic coin-like shape, compacting and bending DNA around the core of the histone octamer (Figure 1D). Although the histone secondary structure is largely preserved (Supplementary Figure S4B), altered histone-DNA and histone-histone contacts lead to subtle changes in its overall architecture (Supplementary Figure S5A). Compared to the *H. sapiens* NCP, the coin face of the *T. brucei* NCP exhibits compression along the horizontal axis (Figure 2A; Supplementary Movie S2), with inwards shifts of histone-DNA interactions at super helical locations (SHL) 2 and 6 (Figure 2B). We did not observe major uneven particle distribution in the cryo-EM map (Supplementary Figure S3H and S3I), suggesting this shape was not due to directional anisotropy. The shift at SHL6 is coincident with a 2-residue insertion in Loop 2 of histone H2A. Although not conserved in sequence, this insertion is found across the *Trypanosoma spp.* (Supplementary Figure S1B, Supplementary Movie S2). The longer length of this loop likely contributes to a ∼5 Å inward tilt of the α2-helix of H2A (Figure 2B). Higher flexibility of this region can also be observed from poorer local resolution (Supplementary Figure S5B).

At SHL2, we observed tighter packing of the α1 helix and loop 1 region (the 'H3 elbow' (4)) of histone H3 to the α2 helix of H4 (Figure 2B, Supplemental Movie 2). This likely occurs due to overall reduction in hydrophobicity and the loss of a bulky aromatic residue (*H. sapiens* ('*Hs*') Phe78 versus *T. brucei* ('*Tb*') Gln75) that shifts the α1-helix of H3 inwards towards H4, increasing DNA bending in order to maintain arginine–phosphate interactions (Supplementary Figure S5C). Combined, these variations in histone architecture result in an oval-shaped NCP with altered DNA binding.

The *T. brucei* NCP has more flexible entry/exit DNA

The ends of the wrapped DNA at the entry/exit site of the *T. brucei* NCP are poorly ordered in the cryo-EM density (Figure 1D). We could reliably model only 126 bp of the 145 bp present, presuming that the missing ends of the DNA are flexibly tethered (Figure 2C). Indeed, during image processing, distinct 3D classes of DNA-ordered states were obtained. Roughly one third of the data indicated a fully wrapped conformation, albeit with a bulged form of DNA, that was previously observed in *Xenopus laevis* NCPs as a precursor to unwrapping (71).The remaining two thirds lacked density around the DNA entry/exit sites (Figure 2D), indicating flexibility of DNA ends.

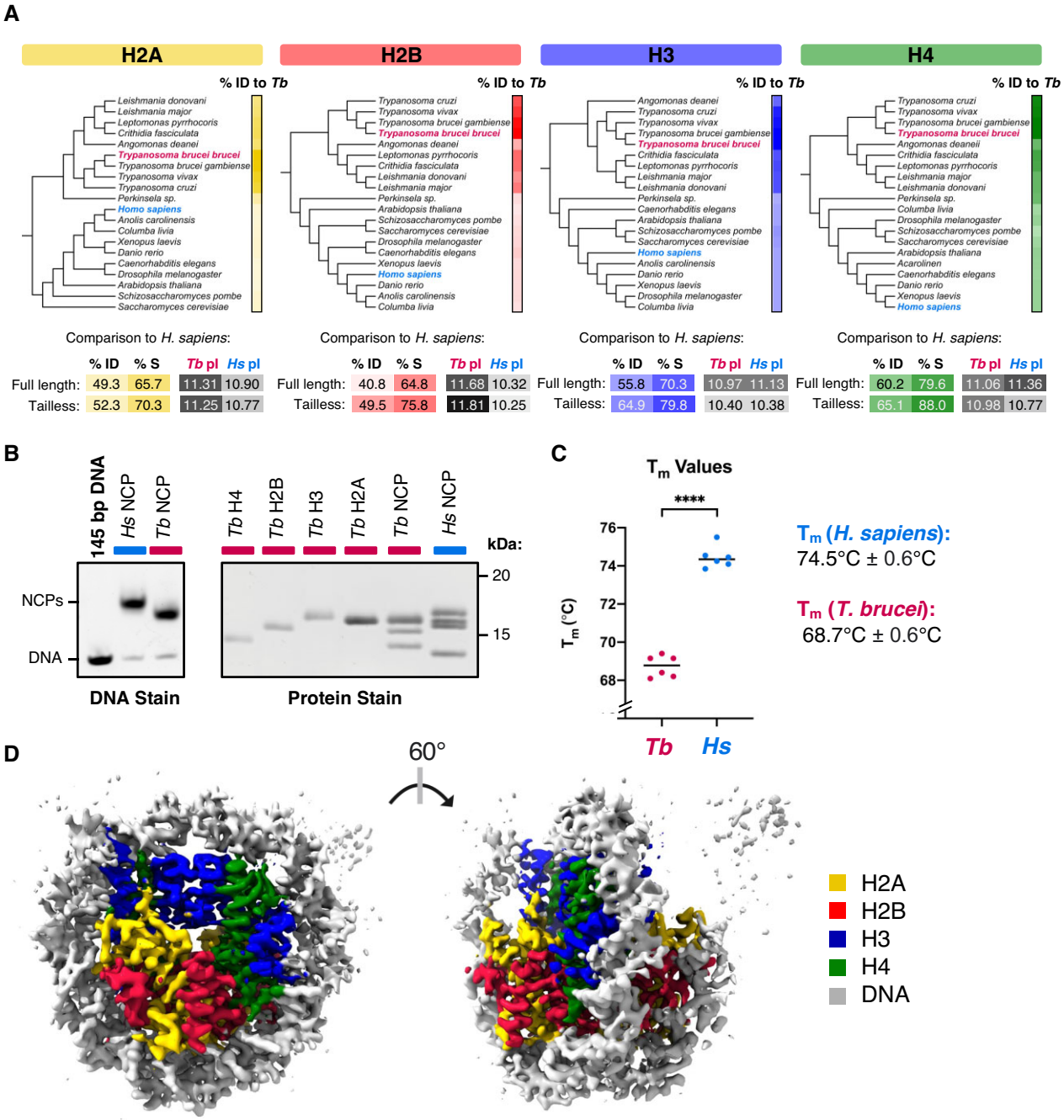


Figure 1. The *T. brucei* nucleosome core particle is evolutionarily divergent. (**A**) Phylogenetic trees constructed from histone sequences of selected model organisms and kinetoplastid species. Scale bars on the right indicate pairwise percentage identity (%ID) to *T. brucei* colored from light to dark (range 40–100%). A comparison of *T. brucei* and *H. sapiens* histone sequences in terms of %ID, percentage similarity (%S), and predicted isoelectric points (pI) is shown below. (**B**) Polyacrylamide gels of *in-vitro* reconstituted *T. brucei* and *H. sapiens* NCPs or component histones in their native (DNA stain) and denatured (protein stain + SDS) states. (**C**) Mean melting temperatures (T_m) of *T. brucei* and *H. sapiens* NCPs from three independent experiments (T_m values were calculated from the first derivative of melting curves shown in Figure 3). (**D**) 3.3 Å cryo-EM density map of the *T. brucei* NCP coloured according to density attributed to histones and DNA.

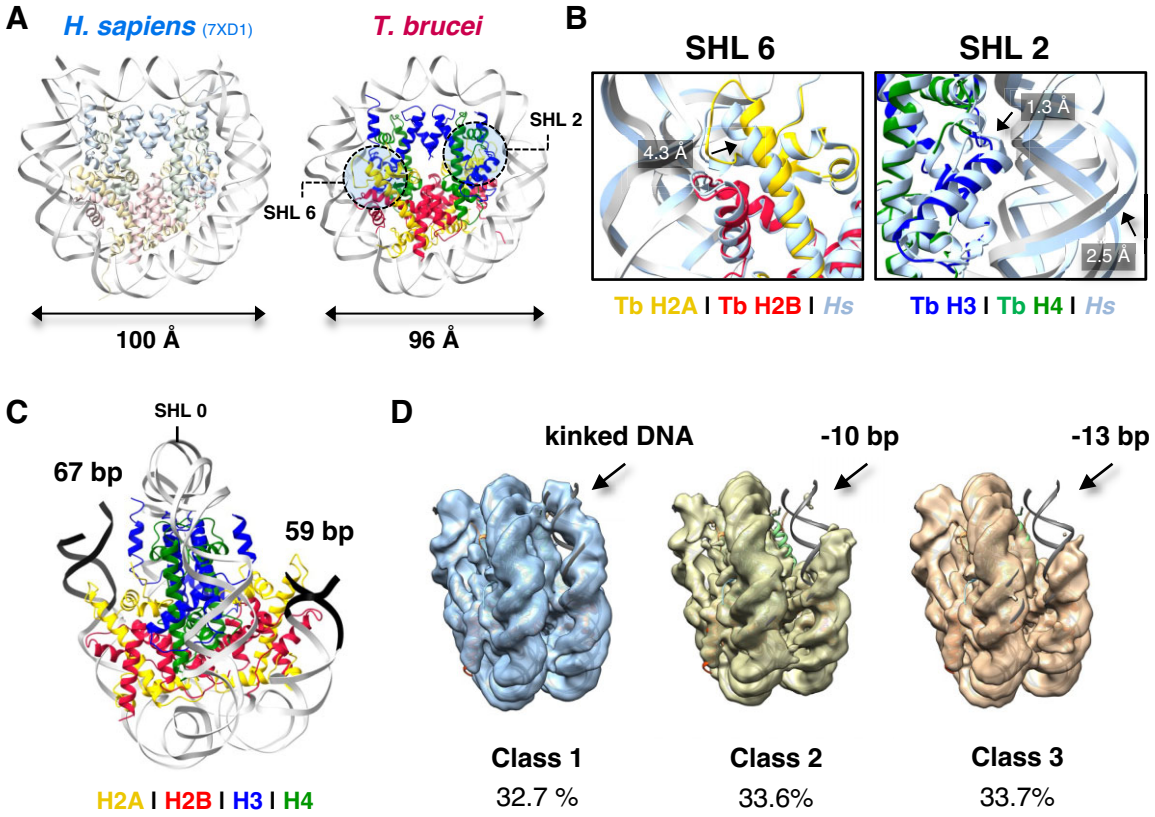


Figure 2. Histone morphology of the *T. brucei* NCP alters its interactions with DNA. (**A**) Estimated diameters of *H. sapiens* (77) and *T. brucei* NCP models from averaged measurements in angstrom (Å) between the phosphate atom of nucleotide –19 and nucleotide –59 in each DNA strand using ChimeraX (46). (**B**) Magnified view of the horizontal compression of the *T. brucei* NCP at SHL6 and SHL2. (**C**) Poorer ordering and asymmetry of DNA ends (colored black) in the *T. brucei* NCP. (**D**) Distinct 3D classes from cryo-EM processing showing different DNA conformations of the *T. brucei* NCP (Class 1 = 45974 particles, Class 2 = 47 192 particles, and Class 3 = 47 429 particles).

The pseudo-symmetry in the *T. brucei* NCP is broken, with one end of the DNA being considerably more disordered. This is likely due to the asymmetry of the Widom 601 sequence (44,72) and is consistent with partial asymmetric unwrapping observed previously (71–75). The inwards compression of DNA observed at SHL2 and 6 and the more oval shape of the *T. brucei* NCP (Figure 2B) may contribute to this splaying of DNA ends. Despite DNA entry/exit site flexibility, we note that the histone-DNA register at the core of the histone octamer is maintained overall, both from comparison to other structures and a hydroxyl radical footprinting assay (Supplementary Figure S5D).

Unsurprisingly, the density for *T. brucei* histone tails was low, arising from a high degree of disorder (76). However, compared to *H. sapiens* and *X. laevis* EM density maps at similar resolutions (71,77), there is poorer ordering for the C-terminal tail of H2A and the N-terminal tail of H3, which engage the final ∼13 bp of straight entry/exit DNA on the NCP (7). *T. brucei* has a number of amino acid substitutions in histones H2A and H3 in these terminal regions (Supplementary Figure S4B). Notably, *Hs* Arg53 is altered to *Tb* Gln50 in the αN helix of histone H3 (Supplementary Figure S5E and S5F). Substitutions at this position were previously shown to be critical for destabilization of entry/exit DNA binding in nucleosomes containing human H3 variants (14,78) DNA end accessibility was also observed in *T. brucei* NCPs using an exonuclease assay (Supplementary Figure S5G). Overall, the oval morphology and poor ordering of H3/H2A tails likely leads to flexibility of entry/exit DNA in the *T. brucei* NCP.

Alterations in histone-histone interfaces in *T. Brucei* NCP lead to instability

Despite the overall conservation of the fold of the *T. brucei* histone octamer, notable changes are present at histone-histone interfaces in the *T. brucei* NCP. Within the H3-H4 tetramer, amino acid substitutions affecting steric packing between H3–H3 and the H3–H4 interface likely alter the stability of the *T. brucei* tetramer and octamer. Electrostatic interactions and the hydrophobic core of the H3–H3 four-helix bundle are disrupted (*Hs* H3-Ala111, *Hs* H3-Ala114 and H3-Leu109 to *Tb* H3-Cys108, *Tb* H3-Ser111 and *Tb* H3-Arg106, respectively) (Figure 3A). Single substitutions in this region have previously been shown to destabilize NCPs (12,79). At the H3-H4 interface, altered hydrophobic interactions arise due to the substitution of *Hs* H3-Phe104 to *Tb* H3-Leu101 (Supplementary Figure S6G). Interestingly, substitutions at this position were also reported in unstable nucleosomes (13,80). Multiple sites that are critical in imparting stability to the H3–H4 tetramer are therefore altered in *T. brucei.*

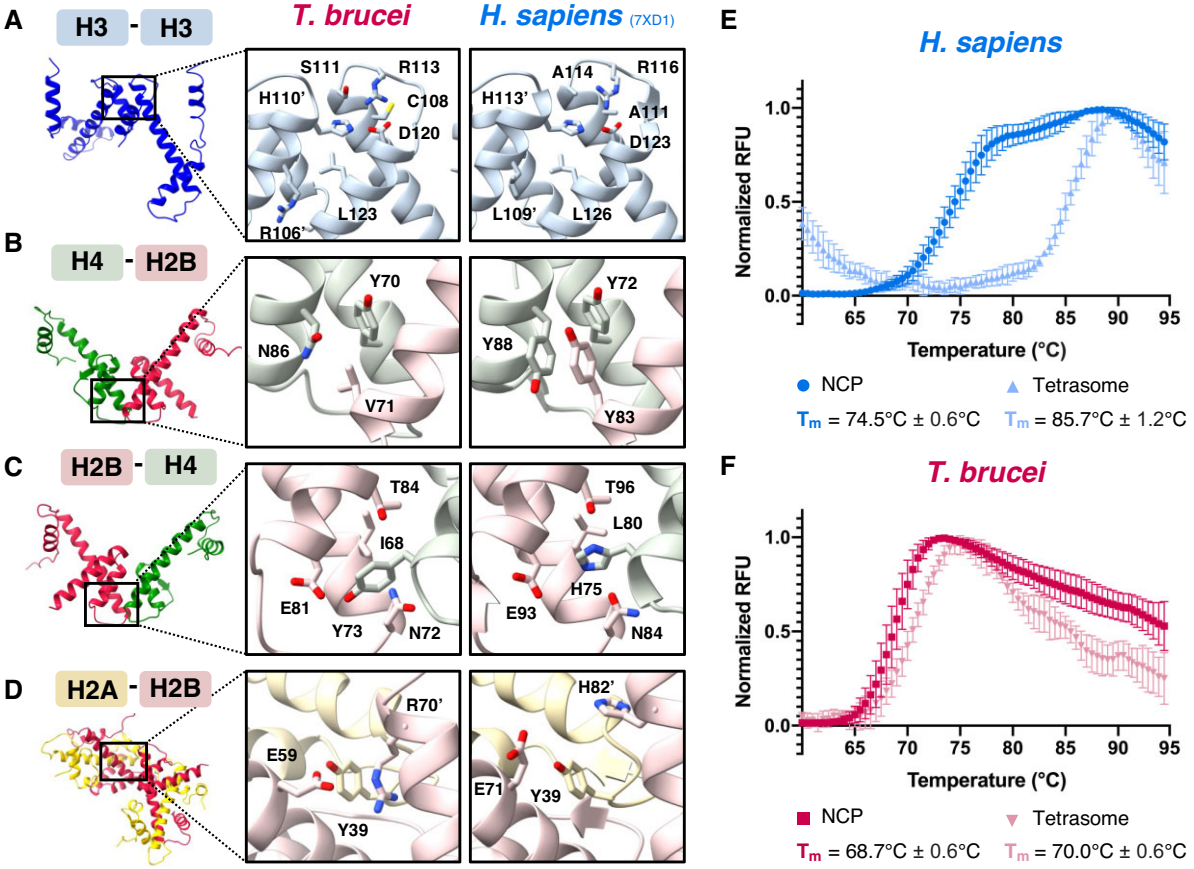


Figure 3. The *T. brucei* NCP is less stable and has a monophasic disassembly pathway. (**A–D**) Magnified views comparing key amino acid interactions at histone-histone interfaces in the *T. brucei* and *H. sapiens* NCPs (77). Apostrophes indicate the second copy of a histone if two histones of the same type are present. (**E**) Biphasic thermal denaturation curves of *H. sapiens* NCPs (darker blue) and monophasic denaturation curves of *H. sapiens* H3–H4 tetrasomes (lighter blue) ($n = 6$). (**F**) Monophasic thermal denaturation curves of *T. brucei* NCPs (darker pink) and *T. brucei* H3-H4 tetrasomes (lighter pink) ($n = 6$). Data points in (E). and (F) are normalized average values ± SD.

Intriguingly, the H3-Cys110 pair that has been implicated in redox sensing (81) is not structurally conserved. Instead, it is replaced by two cysteine pairs that are spatially separated (*Tb* H3-Cys108, *Tb* H3-Cys126).

At the dimer-tetramer interface, the *T. brucei* H2B-H4 four-helix bundle exhibits a loss of aromatic and hydrogen bonding interactions. A π-stacking network mediated by *Hs* H2B-Tyr83, H4-Tyr72 and H4-Tyr88 is absent due to the loss of two of the three tyrosines in *T. brucei* (Figure 3B). A similar loss has been documented in the unstable nucleosome-like structures formed by Melbournevirus histone doublets (82). Additionally, a multivalent hydrogen bonding network centred on H2B-His75 in *H. sapiens* NCPs is disrupted through substitution to H2B Tyr-73 in *T. brucei* (Figure 3C). Conversely, a reinforcement of binding likely occurs at the H2A-H2B dimer-dimer interface, where *Tb* H2A-Tyr39 and *Tb* H2B'-Arg70 likely form a cation-π interaction that is lacking in *H. sapiens* (Figure 3D).

These observations agree well with our attempts to reconstitute hybrid *T. brucei*/ *H. sapiens* histone octamers. Octamers with *Hs* H3 + *Tb* H2A, H2B and H4 could assemble due to the compatibility of *Hs* H3-*Hs* H3 and minor changes at the H3-H4 interface described (Supplementary Figure S2D; Supplementary Figure S6F). NCPs reconstituted with this '*Tb* + *Hs*H3' octamer had slightly higher thermal stability (Supplementary Figure S6H), indicating that the *Tb* H3-H3 interface has lower cohesion. However, more extensively altered histone-histone interfaces prevented the assembly of hybrid octamers containing *Hs* H2A + *Tb* H2B, H3 and H4 (Supplementary Figure S6D) or *Hs* H2A–H2B + *Tb* H3–H4 (Supplementary Figure S6E). Although heterodimerization of *Hs* H2A and *Tb* H2B itself was possible (Supplementary Figure S6D), the formation of a stable H2A-H2B dimer-dimer interface in the context of an assembled octamer was prevented.

The altered protein-protein interactions in *T. brucei* NCPs drive not only lower thermal stability but also result in an altered NCP disassembly pathway (Figure 3E and F). As previously reported (83,84), *H. sapiens* NCPs exhibit biphasic disassembly, whereby the first step involves H2A–H2B disengagement from the H3–H4 tetramer followed by breakdown of the tetramer itself (Figure 3E). Unusually, *T. brucei* NCP disassembly is a single event (Figure 3F). The major drivers are probably the instability of the dimer-tetramer interface and the instability of the tetramer itself, causing both NCP disassembly at a lower temperature and the disassembly of reconstituted *T. brucei* H3-H4 tetrasomes at a similar temperature to *Tb* NCPs (Figure 3E).

Despite their slight increase in stability, the *Tb* + *Hs*H3 NCPs also exhibit monophasic disassembly (Supplementary Figure S6H), indicating inefficient integration of the *Hs*H3 into the altered histone configuration of the *T. brucei* octamer. Interestingly, the majority of the amino acid substitutions that introduce instability in *T. brucei* are conserved in the *Trypanosoma* spp. and would be expected to produce comparable effects in other *Trypanosoma* NCPs (Supplementary Figure 1B).

The *T. brucei* NCP has reduced overall protein-DNA interactions

Alterations in histone–DNA contacts also contribute to the instability of *T. brucei* NCPs. Similarly to *H. sapiens* NCPs, the structure of the histone folds maintains the global alignment of helix dipoles to the DNA phosphate backbone. The total number of histone-DNA hydrogen bonding interactions in the two NCPs is near-equivalent (Figure 4A). However, the distribution of these bonds is altered across the four histone types (Figure 4A) and changes in electrostatic interactions at specific DNA contact points lead to an overall reduction of DNA binding in the *T. brucei* NCP.

At SHL 0 (the dyad), two lysines from histone H3 (*Hs* Lys115) are substituted for glycine (*Tb* Gly112) (Figure 4B). At SHL1.5, a critical arginine in histone H3 (*Hs* Arg63) (7) is substituted for *Tb* Gln60 (Figure 4C). Interestingly, SHL2.5 features a case of histone co-evolution, where histones H2B and H4 in *T. brucei* have a spatially coincident arginine-to-lysine and lysine-to-arginine substitutions respectively (Figure 4D). This prevents steric clash compared to a single substitution and allows for extended interactions with *Tb* H2B-Glu81. However, the new interaction occurs at the expense of the proximity of *Tb* H2B-Lys74 to DNA (Figure 4D).

The N-terminal tail of H2A, which typically straddles the minor groove at SHL4.5 (7,34–35,85) is poorly ordered in *T. brucei* (Figure 4E) and a phosphate-interacting residue, *Hs* Arg11, is replaced by *Tb* Ala11 (Figure 4E and F). The C-terminal helix of histone H2B is displaced and is shorter in length by two residues. The lysine that normally anchors the helix to DNA is therefore lacking (Supplementary Figure S7A). These changes in H2A and H2B are conserved across the kinetoplastid species (Figure 4F, Supplementary Figure S7B).

Cumulatively, the loss of DNA contacts described above suggest weaker DNA binding by the *T. brucei* NCP and this is supported by both lower thermal stability (Figure 1C) and lower resistance to increasing salt conditions compared to the *H. sapiens* NCP (Figure 4G). Interestingly, the *Tb* + *Hs*H3 chimeric NCPs were also more salt labile (Supplementary Figure 7C). This highlights that the highly altered histone configuration in *T. brucei* octamers dominates over integration of new DNA contacts by *Hs*H3.

We also investigated the DNA binding properties of NCPs wrapped with sequences derived from the *T. brucei* genome, namely the 147 bp centromere-associated repeat sequence (86) and the 177 bp minichromosome repeat sequence (87). NCPs could be formed using these native sequences, but *T. brucei* NCPs were harder to reconstitute as a single species compared to *H. sapiens* NCPs (Supplementary Figure S7D). These native sequences were also less well bound than the optimised Widom-601 wrapped NCPs in salt stability assays (Supplementary Figure S7E-G). Furthermore, *T. brucei* NCPs wrapped with native sequences were less stable than *H. sapiens* NCPs, suggesting that the global lower DNA interaction is inherent to *T. brucei* histones rather than due to DNA sequence. Overall, our experiments point towards global reduction of histone-DNA contacts in the *T. brucei* NCP.

A concentrated cluster of positive charge drives DNA binding by H2A–H2B at SHL3.5

While overall DNA-protein contacts are reduced in the *T. brucei* NCP (Figure 4, Supplementary Figure S5D), a unique exception occurs at SHL3.5, where a concentrated cluster of positively charged residues reinforces DNA binding at the dimer-dimer interface of histones H2A and H2B. Compared to *H. sapiens, T. brucei* H2A and H2B have higher calculated net positive charge and this holds true even when the disordered histone tails are discounted (Figure 1A). The structure shows higher electrostatic surface potential at SHL3.5 (Figure 5A) and the interface features net gain of positively-charged side chains from two charge swapping events (*Hs* H2A-Glu41 to *Tb* H2A-Arg41, *Hs* H2B-Glu35 to *Tb* H2B-Arg23), two additional arginine residues (*Tb*-H2B Arg70, *Tb*-H2B Arg75), and a lysine-to-arginine substitution (*Hs* H2A-Lys36 to *Tb*-H2A-Arg36; Figure 5B). The dense cluster of positive charge dramatically alters the overall predicted dipole moment of *T. brucei* and *H. sapiens* histone octamers, where the direction of the dipole is oriented opposite to the dyad at SHL 3.5 (Figure 5C). Interestingly, this may be a kinetoplastid-specific adaptation based on conservation at the sequence level (Supplementary Figure S8A). Furthermore, this interface is similarly charged in models of histone octamers generated using AlphaFold2 (45) from five representative kinetoplastid species, but not other NCP structures (Supplementary Figure S8B).

The observation of locally reinforced DNA interactions by the basic cluster in H2A and H2B correlates well with micrococcal nuclease (MNase) digestion assays. We observed a characteristic digestion band for *T.* brucei NCPs that was observed only weakly in *H. sapiens* NCPs (Figure 5D). We hypothesise this corresponds to the MNase pausing at a stall point, prior to recovering and continuing to digest the DNA. This pattern of digestion was largely independent of the overall rate of MNase digestion (Supplementary Figure S8C).

The same experiment repeated with NCPs reconstituted with tailless *T. brucei* histones did not change the strong stall pattern (Supplementary Figure S8D and S8E). Similarly, the chimeric *Tb* + *Hs*H3 NCPs did not affect the stall (Supplementary Figure S8E). However, the stall was reproduced in *H. sapiens* NCPs by mutating *Hs* H2A-H2B such that the SHL 3.5 contact interface mimics *T. brucei* H2A-H2B (*Hs* H2A R32T K36R E41R and *Hs* H2B E35R K85R R86K S87R) (Figure 5D, Supplementary Figure S8F). Combined, these data demonstrate that the

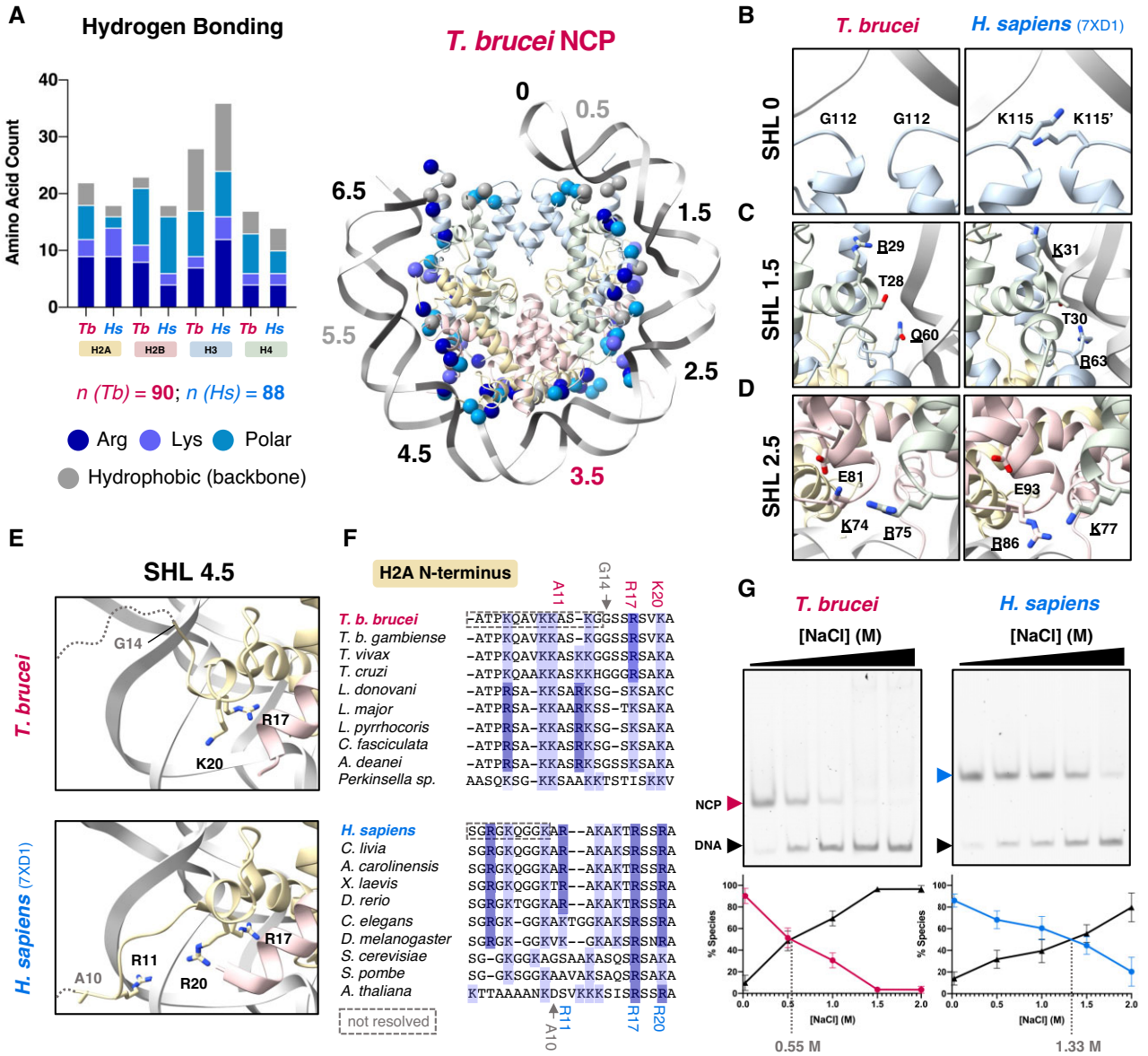


Figure 4. Protein-DNA contacts in the *T. brucei* NCP have a different distribution and result in weaker DNA binding. (**A**) Frequencies of different amino acids forming hydrogen bonds with DNA in the *T. brucei* and *H. sapiens* (PDB: 7XD1) (77) NCPs predicted by ePISA (53) are shown on the left. The total number of interacting residues for each NCP (n) is stated below. On the right, the DNA contacting residues predicted for *T. brucei* are mapped onto the model of the NCP as spheres. SHL annotations are colored according to reduced (black), increased (pink) or similar (grey) interactions with the DNA phosphate backbone. (**B–E**) Reduced protein-DNA interactions at SHL0, 1.5, 2.5 and 4.5 in the *T. brucei* NCP compared to the *H. sapiens* NCP. (**F**) Multiple sequence alignment of the N-terminal tail of histone H2A at SHL 4.5, with DNA-contacting residues from part (E) highlighted. (**G**) Native polyacrylamide gels stained for DNA showing DNA unwrapping in *T. brucei* and *H. sapiens* NCPs after incubation at different NaCl concentrations for 1h on ice. Quantification of the percentage of NCP (*H. sapiens* = blue, *T. brucei* = red) and DNA (black) present in each lane, derived from three biological repeats. The NaCl concentration at which free DNA overtakes wrapped NCPs is indicated in grey.

basic cluster in *T. brucei* H2A–H2B leads to stall of an exonuclease driven by local reinforcement of DNA binding at SHL 3.5.

To approximate DNA length at the 'stall point', we prepared a custom DNA ladder, which showed that the stall point is roughly 10–14 bp shorter than the full-length Widom 601 145 bp DNA in *T. brucei* NCPs (Supplementary Figure S8F). We further mapped the stall point by sequencing both *T. brucei* and *H. sapiens* MNase reaction products (Figure 5E–G). As expected, the digestion occurs asymmetrically; preferentially from the more accessible flexible DNA end (Figure 2C). The stall point corresponds to ~12 bp digestion in *T. brucei* (highest data point in Peak 2, Figure 5G) adjacent to where H2A–H2B first begin to contact DNA (Supplementary Figure S8G). In contrast to *H. sapiens*, further digestion products can barely be detected (Peak 1 in Figure 5F; Supplementary Figure S8G). This is in good agreement with the presence of the highly basic region in the *T. brucei* NCP at SHL 3.5 and may serve as a unique kinetoplastid mechanism to anchor DNA firmly to the base of the nucleosome, despite globally reduced overall DNA binding.

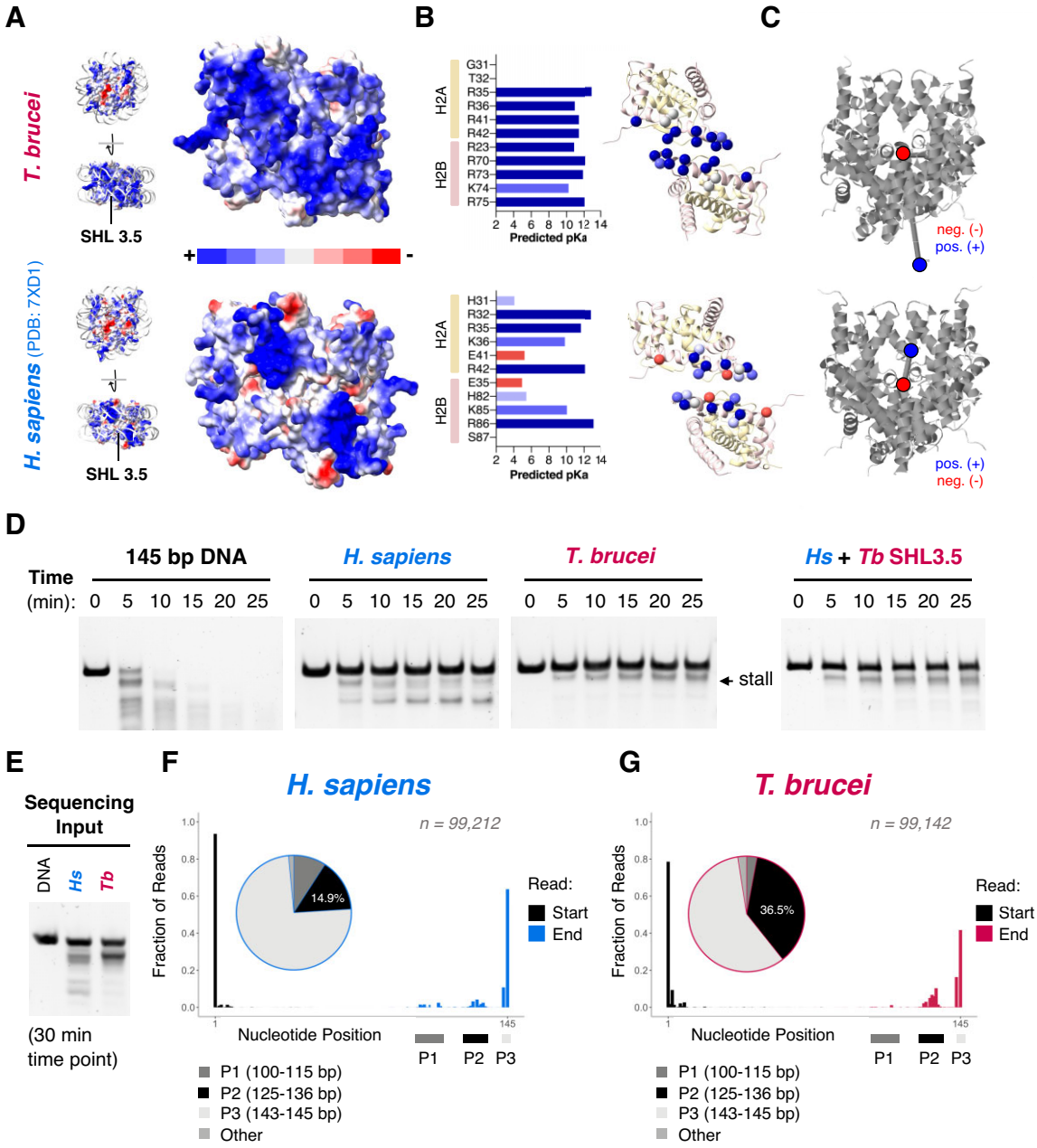


Figure 5. A uniqdue mode of DNA binding at SHL 3.5 in the *T. brucei* NCP. (**A**) Surface electrostatics showing that *T. brucei* histone octamers have a higher density of positive charge at SHL 3.5 compared to *H. sapiens* (77). (**B**) On the left, predicted pK_a values (55) of residues contributing to the electrostatic potential at SHL 3.5 in *T. brucei* and *H. sapiens* NCPs. On the right, the same residues shown as spheres mapped onto the structures of the *T. brucei* and *H. sapiens* H2A-H2B dimers. (**C**) Graphical representation of the predicted overall molecular dipole moments (54) of *T. brucei* (∼1093 Debye) and *H. sapiens* (∼294 Debye) histone octamers (positive end = blue, negative end = red). (**D**) Limited micrococcal nuclease digestion of DNA alone, *H. sapiens* NCPs, *T. brucei* NCPs, and *H. sapiens* NCPs containing mutations in H2A and H2B that mimic the positively charged SHL 3.5 interface in *T. brucei* H2A and H2B ('*H. sapiens* + Tb SHL3.5′ H2A R32T K36R E41R & H2B E35R K85R R86K S87R). (**E**) Scaled up MNase digest of *H. sapiens* and *T. brucei* NCPs used for the sequencing reaction. (**F**) Start and end positions of sequenced *H. sapiens* MNase digestion products from (E) with fractions of reads present in peaks P1, P2 and P3 displayed in a pie chart. (**G**) The same analysis as in (F) repeated for *T. brucei* MNase digestion products.

Altered octamer surface shape and charge in the *T. Brucei* NCP leads to an atypical acidic patch

The acidic patch is a negatively charged region formed between histones H2A and H2B that serves as an interaction platform for many chromatin-associated proteins (4). It has also been implicated in the formation of higher order chromatin structure (7,88). Remarkably, given the high sequence divergence, the negative charge of the canonical acidic patch residues in the *T. brucei* are conserved (Figure 6A). However, the coulombic surface potential of the *T. brucei* acidic patch exhibits stark differences to *H. sapiens* (Figure 6B). The overall shape of the *T. brucei* acidic patch appears narrower and the estimated surface area of the patch is smaller

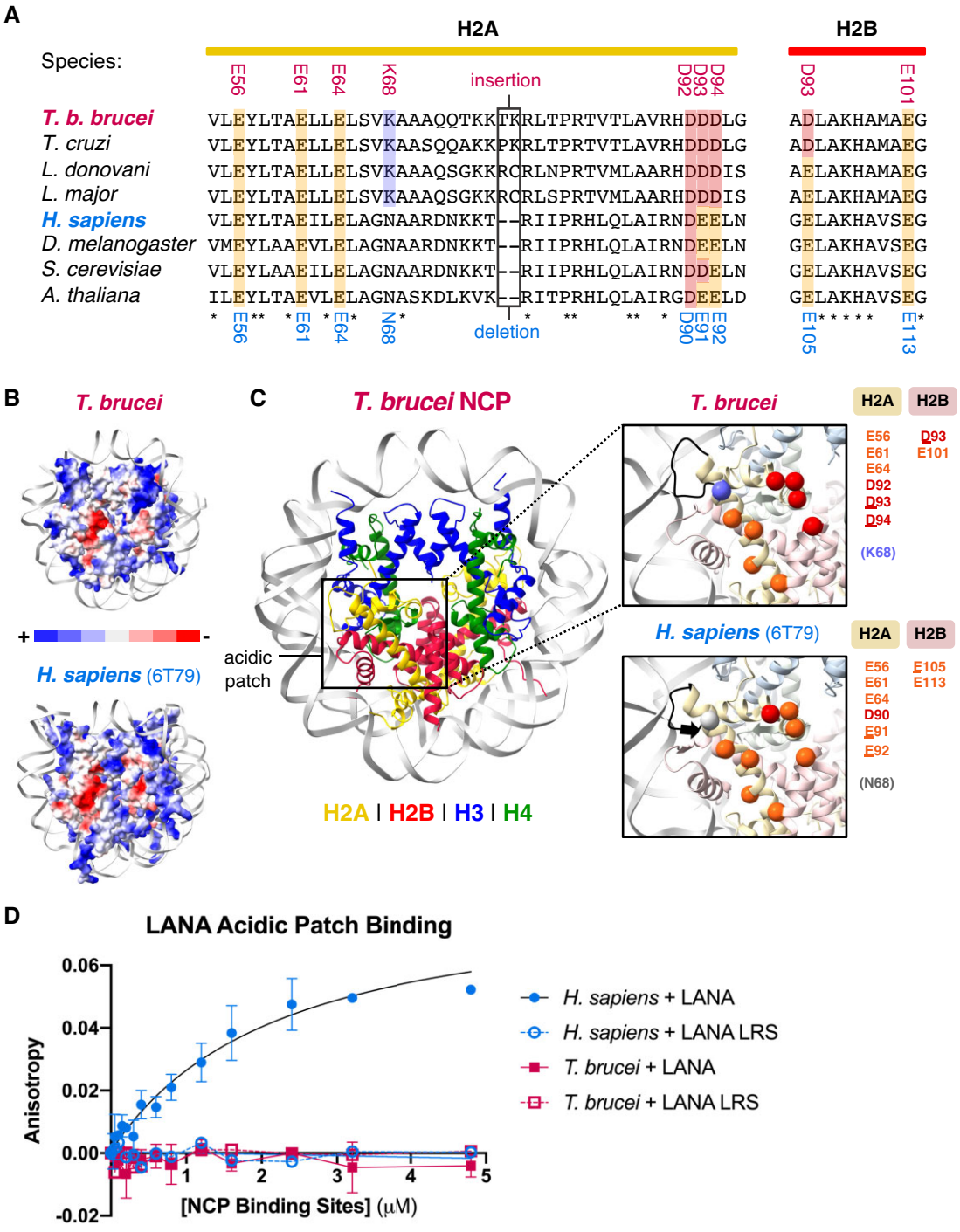


Figure 6. The *T. brucei* acidic patch is highly atypical and refractory to well-characterized binders. (**A**) Multiple sequence alignments of acidic patch regions in H2A and H2B, where residues are highlighted and coloured Glu = orange, Asp = red, Lys = purple, Asn = grey. and the residue numbers for *T. brucei* (above) and *H. sapiens* (below) are indicated. (**B**) The acidic patch region differs when visualizing the electrostatic surface representation of the *T. brucei* (top) and *H. sapiens* (PDB: 6T79 (114), bottom) NCPs. (**C**) The *T. brucei* NCP with a magnified view of the acidic patch and compared to the *H. sapiens* acidic patch below. Acidic patch residues are shown as spheres and coloured using same scheme as (A)**.** The 2-residue insertion in Loop 2 of *T. brucei* H2A is indicated in black. (**D**) Fluorescence polarization assay showing binding of a FITC-tagged LANA peptide and a non-binding mutated LANA peptide (L8A R9A S10A, 'LANA LRS') to *H. sapiens* and *T. brucei* NCPs (K_D for *H. sapiens* NCPs = ~2.38 μM, others ND).

(4498 Å^2 versus 4702 Å^2; Figure 6B). This is likely driven by the two-residue insertion in *Tb* H2A Loop 2 that allows an inwards shift of the α2 helix of H2A at SHL6 (Figure 2B).

A number of other substitutions are present in this region, particularly in histone H2A (Figure 6B). While three of the canonical residues have undergone mild glutamate-to-aspartate substitutions (*Tb* H2A-Asp93, *Tb* H2A-Asp94 and *Tb* H2B-Asp93), a lysine substitution (*Tb* H2A-Lys68 from *Hs* H2A-Asn68) is in close proximity to the patch (Figure 6C). This substitution is conserved across the kinetoplastids (Supplementary Figure S9A) Interestingly, mutation of *Hs* H2A-Asn68 was recently shown to reduce binding to multiple chromatin-associated proteins in a proteome-wide screen (89). Both the surface charge and shape alterations are also conserved in our predicted models of five other kinetoplastid histone octamers (Supplementary Figure S9B).

These differences in the *T. brucei* acidic patch likely have direct functional consequences for chromatin reading. A prototypical acidic patch binder, the Kaposi's Sarcoma Herpesvirus Latency Associated Nuclear Antigen (LANA) peptide (68), binds well to the *H. sapiens* NCP but showed no detectable interaction to the *T. brucei* NCP in a fluorescence polarization assay (Figure 6D). Similar results were obtained for an acidic patch binder with an altered binding mode (Supplementary Figure 9C), the Prototype Foamy Virus GAG peptide (PFV-GAG) (36), using both fluorescence polarization (Supplementary Figure 9D) and electrophoretic mobility shift assays (Supplementary Figure S9E). Discrete mutations to individual components of the *T. brucei* acidic patch were not sufficient to rescue the binding of the PFV-GAG and LANA peptides (Supplementary Figure S9A, S9E and S9F). These results reveal extensive divergence of this common protein interacting hub and indicate that the kinetoplastids likely have altered chromatin-protein and higher order chromatin interactions.

DISCUSSION

In this study, we present the cryo-EM structure and biochemical characterisation of the nucleosome core particle from the parasitic protist *T. brucei*. Despite extensive divergence of *T. brucei* histone sequences (Figure 1), the structure exhibits remarkable conservation of overall histone fold architecture. This is consistent with structures from other divergent organisms such as the parasite *Giardia lamblia* (90), the archaeon *Methanothermus fervidus* (91), or the *Marseilleviridae* (82,92) giant viruses. However, subtle differences in histone packing and specific histone amino acid substitutions give rise to properties unique to the *T. brucei* NCP. The *T. brucei* NCP has a compressed shape (Figure 2, Supplementary Movie S2) and four out of seven DNA contact points at half-integral SHLs have lost critical histone-DNA interactions, leading to flexible DNA ends and low stability *in vitro* (Figure 2–4). A striking compensation mechanism occurs via H2A–H2B dimers at SHL3.5, whereby DNA binding is increased by a concentrated cluster of basic residues (Figure 5). Furthermore, kinetoplastid-specific alterations lead to an altered topology and charge of common protein interaction interface, the acidic patch (Figure 6). By comparing the molecular differences in nucleosomes between trypanosomes and conventional eukaryotic systems, we can hope to understand both the evolutionary constraints and diversity that directs DNA-associated mechanisms.

Early studies of trypanosome chromatin showed that *T. brucei* chromatin is less stable (93) and has reduced higher order compaction than is observed in model eukaryotes (21,94,95). However, at the sequence level, the *T. brucei* genome is not unusual in base composition nor less enriched for nucleosome positioning features (96,97). Our results indicate that chromatin instability in *T. brucei* is inherent at the mononucleosome level via weakened histone–histone interfaces, histone-DNA contacts, and flexible entry/exit DNA (Figure 2–4). Indeed, NCPs could be wrapped using native *T. brucei* DNA sequences, but these were also less stable compared to *H. sapiens* NCPs wrapped using the same sequences (Supplementary Figure 7).

Chromatin arrays constructed from other nucleosomes with increased DNA end flexibility (80,90,98) were previously found to favour more open chromatin conformations, particularly due to alterations in inter-nucleosomal DNA path (98). Future experiments on chromatin arrays would help explore if entry/exit DNA flexibility described here could explain the more open chromatin in *T. brucei*.

The end flexibility we observed was reminiscent of nucleosomes reconstituted with the centromeric H3 variant CENP-A (67,72). Interestingly, this variant is missing from *T. brucei* and how centromere regions are specified is currently unclear (99). Six of the *T. brucei* megachromosomes contain a 147 bp centromere-associated repeat sequence (86), which we show can be wrapped into NCPs (Supplementary Figure 7D). Investigating epigenetic inheritance and the role of chromatin in centromere-associated processes will be of interest in the future.

However, despite overall lower DNA binding of the *T. brucei* NCP, histones H2A and H2B in *T. brucei* were previously shown to be bound to DNA more persistently than H3 and H4 (100). This is consistent with our finding that H2A-H2B reinforce DNA binding at SHL3.5 via a large positive dipole (Figure 5). The significance of tight DNA binding by H2A–H2B requires further investigation but we speculate that it may affect a variety of chromatin-based processes. For example, RNA polymerase II transcription is known to be stalled at defined points corresponding to DNA-histone contacts while transcribing a nucleosome template (101–104). Transcription in *T. brucei* is expected to be highly atypical (105) and future studies could reveal whether the altered DNA binding properties of their nucleosomes affect this and other processes.

Recent work has shown the complexity of histone post-translational marks in *T. brucei* (24,26,27) and *T. cruzi* (106,107). Although most of the N- and C-terminal histone tails are not resolved in our structure, we could map 47 reported histone marks onto our unmodified structure (24,26) (Supplementary Figure S10A). Intriguingly, the kinetoplastid-specific H2A-Lys68, which alters the local environment of the acidic patch (Figure 6B), has been reported to be trimethylated (Supplementary Figure S10A). Conceivably, this and other modifications may act

as switches to alter chromatin binding properties. We also note a cluster of phosphorylation/acetylation sites at the N-terminal end of H2A (Supplementary Figure S10A) that could serve to further destabilise DNA binding at SHL 4.5 (Figure 4E). Beyond histone post translational modifications, histone variants in *T. brucei* play a key role in modulating transcription (22–24) and controlling the parasite's variant surface glycoprotein immune evasion system (19). Further work comparing the canonical structure of *T. brucei* nucleosomes to variants would be of great interest to help explain the biology behind antigenic variation and transcription initiation/termination.

Of the four common interaction interfaces commonly co-opted by chromatin-binding proteins (4), all display both sequence and structural variation in *T. brucei*. For example, the acidic patch in *T. brucei* has altered topology and is refractory to known binders (Figure 6). Interestingly, the narrow topology of the patch accompanied by an insertion in H2A in *T. brucei* is inverse to the widening of the patch with an insertion in H2B in the nucleosome structure of the parasite *G. lamblia* (90) (Supplementary Figure S9A). The divergence of this interface suggests that trypanosome chromatin interactions may also have divergent properties. Since the acidic patch has been identified as a common interaction site in other species (4), we expect that other identified factors in *T. brucei* involved in regulating histone post-translational modifications (27), chromatin remodelling (28,29), antigenic variation (30), or yet unidentified pathways may use this mechanism.

However, a mechanistic understanding for chromatin interactions in *T. brucei* is largely missing and our search for acidic-patch binders *in T. brucei* was challenging due to low conservation or poor annotation of potential homologs. Despite this, promising candidates could include *T. brucei* Dot1A and Dot1B, homologs of the Dot1 histone methyltransferase in higher eukaryotes (108–110). The catalytic fold of *T. brucei* Dot1A/B seems to be conserved (31) when modelled using AlphaFold2 (45) (Supplementary Figure S10B), but the loop that engages the acidic patch in human Dot1L (110) (Supplementary Figure S10C) differs in both sequence and predicted structure in *T. brucei* (Supplementary Figure S10D and E). This suggests that the local binding interactions to chromatin by *T. brucei* Dot1A/B may differ. It will be a fascinating avenue of study to probe the effects of divergence on chromatin recognition in kinetoplastid parasites.

Remarkably, our extensive phylogenetic analysis and modelling revealed that a majority of our findings are conserved within the kinetoplastids, including known pathogens (Supplementary Figures S1A, S8A, B, S9A, S8B). For example, conservation of the altered acidic patch includes clinically relevant targets from the *Trypanosoma* and *Leishmania* species (Supplementary Figure S9A, B). This opens a possible therapeutic avenue for targeting the acidic patch (111,112) allowing specific targeting of kinetoplastids over their human or animal hosts. A global chromatin disruption mechanism would have high utility for combating diseases such as animal trypanosomiasis where the genetic diversity of *Trypanosoma* species has hindered drug development (17) and drug resistance is a current challenge (16,113).

DATA AVAILABILITY

The cryo-EM density map and associated meta data for the *T. brucei* NCP have been deposited at the Electron Microscopy Data Bank under accession number EMD-16777. Raw micrographs have been uploaded to EMPIAR-11454. The atomic coordinates of the *T. brucei* NCP have been deposited in the Protein Data Bank under accession number 8COM. The MNase sequencing results have been deposited to the Gene Expression Omnibus NCBI database under accession number GSE226029.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Robin Allshire, David Horn, Christopher Wood, Elizabeth Blackburn, and members of the Wilson lab for helpful discussions and feedback. We thank Kurt Davey for gift of the alpha satellite repeat plasmid and Christian Janzen for gift of the *T. brucei* histone plasmids. Human histones were supplied from Addgene from the Landry lab. We thank Logan Mackay in SIRCAMS, School of Chemistry, University of Edinburgh for mass spec analysis. We also thank the Technical Services Sequencing Facility at The Institute of Genetics and Cancer, University of Edinburgh for processing hydroxyl radical footprinting samples. We are very grateful to Yuriy Chaban at Diamond Light Source for access and support of the Cryo-EM facilities at eBIC. We are also grateful to Nathan Cowieson at Diamond Light Source for access and support.
Author contributions: M.D.W. conceived the study and supervised the project. G.D. and M.D.W. designed the experiments (unless otherwise stated), analyzed the data, and wrote the manuscript, with input from the other authors. M.R.D.T., R.C., G.D., M.D.W. and H.W. purified protein and DNA components. Cryo-EM image processing was performed by M.D.W., G.S. and M.W.T. Model building was performed by M.D.W., G.D. and H.B. J.W. performed DNA footprinting assays. Biochemical assays were performed by M.D.W., G.D. and H.W. Phylogenetic analysis and computational protein structure analysis was performed by G.D. MNase sequencing data were processed by S.W. and G.D.

FUNDING

M.D.W.'s work is supported by the Wellcome Trust [210493]; Medical Research Council [T029471/1]; University of Edinburgh; G.D.'s work is supported by BBSRC EASTBIO [BB/M010996/1]; initial grid screening for cryo-EM was performed in the cryo-EM facility in School of Biological Sciences at the University of Edinburgh, which was set up with funding from the Wellcome Trust [087658/Z/08/Z]; SULSA; the cryo-EM session (proposal EM-BI24557) at the UK national electron bio-imaging centre (eBIC) was funded by the Wellcome Trust, MRC and BBSRC; the SEC-SAXS was supported at the B21 beamline (rapid access proposal SM27169); Edinburgh Protein Production Facility (EPPF) and the WCB Bioinformatics Core

Facility, both of which receive funding from a core grant [203149] to the Wellcome Centre for Cell Biology at the University of Edinburgh. Funding for open access charge: Wellcome Trust.
Conflict of interest statement. None declared.

REFERENCES

1. Noll,M. (1974) Subunit structure of chromatin. *Nature*, **251**, 249–251.
2. Bednar,J., Horowitz,R.A., Grigoryev,S.A., Carruthers,L.M., Hansen,J.C., Koster,A.J. and Woodcock,C.L. (1998) Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 14173–14178.
3. Kornberg,R.D. and Lorch,Y. (2020) Primary role of the nucleosome. *Mol. Cell*, **79**, 371–375.
4. McGinty,R.K. and Tan,S. (2021) Principles of nucleosome recognition by chromatin factors and enzymes. *Curr. Opin. Struct. Biol.*, **71**, 16–26.
5. Malik,H.S. and Henikoff,S. (2003) Phylogenomics of the nucleosome. *Nat. Struct. Mol. Biol.*, **10**, 882–891.
6. Grau-Bové,X., Navarrete,C., Chiva,C., Pribasnig,T., Antó,M., Torruella,G., Galindo,L.J., Lang,B.F., Moreira,D., López-Garcia,P. *et al.* (2022) A phylogenetic and proteomic reconstruction of eukaryotic chromatin evolution. *Nat. Ecol. Evol.*, **6**, 1007–1023.
7. Luger,K., Mäder,A.W., Richmond,R.K., Sargent,D.F. and Richmond,T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**, 251–260.
8. Tsunaka,Y. (2005) Alteration of the nucleosomal DNA path in the crystal structure of a human nucleosome core particle. *Nucleic Acids Res.*, **33**, 3424–3434.
9. Clapier,C.R., Chakravarthy,S., Petosa,C., Fernández-Tornero,C., Luger,K. and Müller,C.W. (2008) Structure of theDrosophila nucleosome core particle highlights evolutionary constraints on the H2A-H2B histone dimer. *Proteins*, **71**, 1–7.
10. White,C.L., Suto,R.K. and Luger,K. (2001) Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J.*, **20**, 5207–5218.
11. Fukushima,Y., Hatazawa,S., Hirai,S., Kujirai,T., Ehara,H., Sekine,S.-i., Takizawa,Y. and Kurumizaka,H. (2022) Structural and biochemical analyses of the nucleosome containing *Komagataella pastoris*histones. *J. Biochem.*, **172**, 79–88.
12. Tachiwana,H., Kagawa,W., Osakabe,A., Kawaguchi,K., Shiga,T., Hayashi-Takanaka,Y., Kimura,H. and Kurumizaka,H. (2010) Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 10454–10459.
13. Urahama,T., Harada,A., Maehara,K., Horikoshi,N., Sato,K., Sato,Y., Shiraishi,K., Sugino,N., Osakabe,A., Tachiwana,H. *et al.* (2016) Histone H3.5 forms an unstable nucleosome and accumulates around transcription start sites in human testis. *Epigenet. Chromatin*, **9**, 2.
14. Hirai,S., Tomimatsu,K., Miyawaki-Kuwakado,A., Takizawa,Y., Komatsu,T., Tachibana,T., Fukushima,Y., Takeda,Y., Negishi,L., Kujirai,T. *et al.* (2022) Unusual nucleosome formation and transcriptome influence by the histone H3mm18 variant. *Nucleic Acids Res.*, **50**, 72–91.
15. Lukeš,J., Skalický,T., Týč,J., Votýpka,J. and Yurchenko,V. (2014) Evolution of parasitism in kinetoplastid flagellates. *Mol. Biochem. Parasitol.*, **195**, 115–122.
16. Richards,S., Morrison,L.J., Torr,S.J., Barrett,M.P., Manangwa,O., Mramba,F. and Auty,H. (2021) Pharma to farmer: field challenges of optimizing trypanocide use in African animal trypanosomiasis. *Trends Parasitol.*, **37**, 831–843.
17. Morrison,L.J., Steketee,P.C., Tettey,M.D. and Matthews,K.R. (2023) Pathogenicity and virulence of African trypanosomes: from laboratory models to clinically relevant hosts. *Virulence*, **14**, 2150445.
18. De Rycker,M., Wyllie,S., Horn,D., Read,K.D. and Gilbert,I.H. (2023) Anti-trypanosomatid drug discovery: progress and challenges. *Nat. Rev. Microbiol.*, **21**, 35–50.
19. Müller,L.S.M., Cosentino,R.O., Förstner,K.U., Guizetti,J., Wedel,C., Kaplan,N., Janzen,C.J., Arampatzi,P., Vogel,J., Steinbiss,S. *et al.* (2018) Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature*, **563**, 121–125.
20. Berriman,M., Ghedin,E., Hertz-Fowler,C., Blandin,G., Renauld,H., Bartholomeu,D.C., Lennard,N.J., Caler,E., Hamlin,N.E., Haas,B. *et al.* (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science*, **309**, 416–422.
21. Yamin,K., Bijlani,S., Berman,J., Soni,A., Shlomai,J., Buragohain,B.M., Werbner,M., Gal-Tanamy,M., Matityahu,A. and Onn,I. (2022) Fold-change of chromatin condensation in yeast is a conserved property. *Sci. Rep.*, **12**, 17393.
22. Siegel,T.N., Hekstra,D.R., Kemp,L.E., Figueiredo,L.M., Lowell,J.E., Fenyo,D., Wang,X., Dewell,S. and Cross,G.A.M. (2009) Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes Dev.*, **23**, 1063–1076.
23. Schulz,D., Zaringhalam,M., Papavasiliou,F.N. and Kim,H.-S. (2016) Base J and H3.V regulate transcriptional termination in *Trypanosoma brucei*. *PLoS Genet.*, **12**, e1005762.
24. Kraus,A.J., Vanselow,J.T., Lamer,S., Brink,B.G., Schlosser,A. and Siegel,T.N. (2020) Distinct roles for H4 and H2A.Z acetylation in RNA transcription in African trypanosomes. *Nat. Commun.*, **11**, 1498.
25. Alsford,S. and Horn,D. (2004) Trypanosomatid histones: trypanosomatid histones. *Mol. Microbiol.*, **53**, 365–372.
26. Maree,J.P., Tvardovskiy,A., Ravnsborg,T., Jensen,O.N., Rudenko,G. and Patterton,H.-G. (2022) *Trypanosoma brucei*histones are heavily modified with combinatorial post-translational modifications and mark Pol II transcription start regions with hyperacetylated H2A. *Nucleic Acids Res.*, **50**, 9705–9723.
27. Staneva,D.P., Carloni,R., Auchynnikava,T., Tong,P., Rappsilber,J., Jeyaprakash,A.A., Matthews,K.R. and Allshire,R.C. (2021) A systematic analysis of *Trypanosoma brucei*chromatin factors identifies novel protein interaction networks associated with sites of transcription initiation and termination. *Genome Res.*, **31**, 2138–2154.
28. Vellmer,T., Hartleb,L., Fradera Sola,A., Kramer,S., Meyer-Natus,E., Butter,F. and Janzen,C.J. (2022) A novel SNF2 ATPase complex in *Trypanosoma brucei* with a role in H2A.Z-mediated chromatin remodelling. *PLoS Pathog.*, **18**, e1010514.
29. Stanne,T., Narayanan,M.S., Ridewood,S., Ling,A., Witmer,K., Kushwaha,M., Wiesler,S., Wickstead,B., Wood,J. and Rudenko,G. (2015) Identification of the ISWI chromatin remodeling complex of the early branching eukaryote *Trypanosoma brucei*. *J. Biol. Chem.*, **290**, 26954–26967.
30. Faria,J., Glover,L., Hutchinson,S., Boehm,C., Field,M.C. and Horn,D. (2019) Monoallelic expression and epigenetic inheritance sustained by a *Trypanosoma brucei* variant surface glycoprotein exclusion complex. *Nat. Commun.*, **10**, 3023.
31. Dindar,G., Anger,A.M., Mehlhorn,C., Hake,S.B. and Janzen,C.J. (2014) Structure-guided mutational analysis reveals the functional requirements for product specificity of DOT1 enzymes. *Nat. Commun.*, **5**, 5313.
32. Iwasaki,W., Miya,Y., Horikoshi,N., Osakabe,A., Taguchi,H., Tachiwana,H., Shibata,T., Kagawa,W. and Kurumizaka,H. (2013) Contribution of histone N-terminal tails to the structure and stability of nucleosomes. *FEBS Open Bio*, **3**, 363–369.
33. Dyer,P.N., Edayathumangalam,R.S., White,C.L., Bao,Y., Chakravarthy,S., Muthurajan,U.M. and Luger,K. (2003) In: *Methods in Enzymology*. Elsevier, Vol. **375**, pp. 23–44.
34. Wilson,M.D., Benlekbir,S., Fradet-Turcotte,A., Sherker,A., Julien,J.-P., McEwan,A., Noordermeer,S.M., Sicheri,F., Rubinstein,J.L. and Durocher,D. (2016) The structural basis of modified nucleosome recognition by 53BP1. *Nature*, **536**, 100–103.
35. Wilson,M.D., Renault,L., Maskell,D.P., Ghoneim,M., Pye,V.E., Nans,A., Rueda,D.S., Cherepanov,P. and Costa,A. (2019) Retroviral integration into nucleosomes through DNA looping and sliding along the histone octamer. *Nat. Commun.*, **10**, 4189.
36. Lesbats,P., Serrao,E., Maskell,D.P., Pye,V.E., O'Reilly,N., Lindemann,D., Engelman,A.N. and Cherepanov,P. (2017)

Structural basis for spumavirus GAG tethering to chromatin. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 5509–5514.
37. Maskell,D.P., Renault,L., Serrao,E., Lesbats,P., Matadeen,R., Hare,S., Lindemann,D., Engelman,A.N., Costa,A. and Cherepanov,P. (2015) Structural basis for retroviral integration into nucleosomes. *Nature*, **523**, 366–369.
38. Leong,S.L., Lynch,E.M., Zou,J., Tay,Y.D., Borek,W.E., Tuijtel,M.W., Rappsilber,J. and Sawin,K.E. (2019) Reconstitution of microtubule nucleation in vitro reveals novel roles for Mzt1. *Curr. Biol.*, **29**, 2199–2207.
39. Mastronarde,D.N. (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.*, **152**, 36–51.
40. Punjani,A., Rubinstein,J.L., Fleet,D.J. and Brubaker,M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*, **14**, 290–296.
41. Punjani,A., Zhang,H. and Fleet,D.J. (2020) Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods*, **17**, 1214–1221.
42. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera?A visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
43. Tan,Y.Z., Baldwin,P.R., Davis,J.H., Williamson,J.R., Potter,C.S., Carragher,B. and Lyumkis,D. (2017) Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods*, **14**, 793–796.
44. Vasudevan,D., Chua,E.Y.D. and Davey,C.A. (2010) Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. *J. Mol. Biol.*, **403**, 1–10.
45. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.
46. Pettersen,E.F., Goddard,T.D., Huang,C.C., Meng,E.C., Couch,G.S., Croll,T.I., Morris,J.H. and Ferrin,T.E. (2021) UCSF CHIMERAX: structure visualization for researchers, educators, and developers. *Protein Sci.*, **30**, 70–82.
47. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of *Coot*. *Acta. Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
48. Croll,T.I. (2018) *ISOLDE*: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.*, **74**, 519–530.
49. Afonine,P.V., Poon,B.K., Read,R.J., Sobolev,O.V., Terwilliger,T.C., Urzhumtsev,A. and Adams,P.D. (2018) Real-space refinement in *PHENIX*for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.*, **74**, 531–544.
50. Chen,V.B., Arendall,W.B., Headd,J.J., Keedy,D.A., Immormino,R.M., Kapral,G.J., Murray,L.W., Richardson,J.S. and Richardson,D.C. (2010) *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta. Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.
51. Barad,B.A., Echols,N., Wang,R.Y.-R., Cheng,Y., DiMaio,F., Adams,P.D. and Fraser,J.S. (2015) EMRinger: side chain–directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods*, **12**, 943–946.
52. Cowieson,N.P., Edwards-Gayle,C.J.C., Inoue,K., Khunti,N.S., Doutch,J., Williams,E., Daniels,S., Preece,G., Krumpa,N.A., Sutter,J.P. *et al.* (2020) Beamline B21: high-throughput small-angle X-ray scattering at Diamond Light Source. *J. Synchrotron Radiat.*, **27**, 1438–1446.
53. Krissinel,E. and Henrick,K. (2007) Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.*, **372**, 774–797.
54. Felder,C.E., Prilusky,J., Silman,I. and Sussman,J.L. (2007) A server and database for dipole moments of proteins. *Nucleic Acids Res.*, **35**, W512–W521.
55. Olsson,M.H.M., Søndergaard,C.R., Rostkowski,M. and Jensen,J.H. (2011) PROPKA3: consistent treatment of internal and surface residues in empirical pK_a predictions. *J. Chem. Theory Comput.*, **7**, 525–537.
56. Amos,B., Aurrecoechea,C., Barba,M., Barreto,A., Basenko,E.Y., Bażant,W., Belnap,R., Blevins,A.S., Böhme,U., Brestelli,J. *et al.* (2022) VEuPathDB: the eukaryotic pathogen, vector and host bioinformatics resource center. *Nucleic Acids Res.*, **50**, D898–D911.
57. Sayers,E.W., Bolton,E.E., Brister,J.R., Canese,K., Chan,J., Comeau,D.C., Connor,R., Funk,K., Kelly,C., Kim,S. *et al.* (2022) Database resources of the national center for biotechnology information. *Nucleic Acids Res.*, **50**, D20–D26.
58. Katoh,K. and Standley,D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, **30**, 772–780.
59. Waterhouse,A.M., Procter,J.B., Martin,D.M.A., Clamp,M. and Barton,G.J. (2009) Jalview Version 2 – a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.
60. Nguyen,L.-T., Schmidt,H.A., von Haeseler,A. and Minh,B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.*, **32**, 268–274.
61. Letunic,I. and Bork,P. (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, **23**, 127–128.
62. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.
63. Needleman,S.B. and Wunsch,C.D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.*, **48**, 443–453.
64. Gasteiger,E., Hoogland,C., Gattiker,A., Duvaud,S.e., Wilkins,M.R., Appel,R.D. and Bairoch,A. (2005) In: Walker,J.M. (ed). *The Proteomics Protocols Handbook*. Humana Press, Totowa, NJ, pp. 571–607.
65. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.*, **17**, 10.
66. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
67. Tachiwana,H., Kagawa,W., Shiga,T., Osakabe,A., Miya,Y., Saito,K., Hayashi-Takanaka,Y., Oda,T., Sato,M., Park,S.-Y. *et al.* (2011) Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature*, **476**, 232–235.
68. Barbera,A.J., Chodaparambil,J.V., Kelley-Clarke,B., Joukov,V., Walter,J.C., Luger,K. and Kaye,K.M. (2006) The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science*, **311**, 856–861.
69. Beauchemin,C., Moerke,N.J., Faloon,P. and Kaye,K.M. (2014) Assay development and high-throughput screening for inhibitors of Kaposi's sarcoma–associated herpesvirus N-terminal latency-associated nuclear antigen binding to nucleosomes. *SLAS Discovery*, **19**, 947–958.
70. Stevens,J.R. (2008) Kinetoplastid phylogenetics, with special reference to the evolution of parasitic trypanosomes. *Parasite*, **15**, 226–232.
71. Bilokapic,S., Strauss,M. and Halic,M. (2018) Histone octamer rearranges to adapt to DNA unwrapping. *Nat. Struct. Mol. Biol.*, **25**, 101–108.
72. Boopathi,R., Danev,R., Khoshouei,M., Kale,S., Nahata,S., Ramos,L., Angelov,D., Dimitrov,S., Hamiche,A., Petosa,C. *et al.* (2020) Phase-plate cryo-EM structure of the Widom 601 CENP-A nucleosome core particle reveals differential flexibility of the DNA ends. *Nucleic Acids Res.*, **48**, 5735–5748.
73. Chen,Y., Tokuda,J.M., Topping,T., Sutton,J.L., Meisburger,S.P., Pabit,S.A., Gloss,L.M. and Pollack,L. (2014) Revealing transient structures of nucleosomes as DNA unwinds. *Nucleic Acids Res.*, **42**, 8767–8776.
74. Ngo,T.T.M., Zhang,Q., Zhou,R., Yodh,J.G. and Ha,T. (2015) Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. *Cell*, **160**, 1135–1144.
75. North,J.A., Shimko,J.C., Javaid,S., Mooney,A.M., Shoffner,M.A., Rose,S.D., Bundschuh,R., Fishel,R., Ottesen,J.J. and Poirier,M.G. (2012) Regulation of the nucleosome unwrapping rate controls DNA accessibility. *Nucleic Acids Res.*, **40**, 10215–10227.
76. Ghoneim,M., Fuchs,H.A. and Musselman,C.A. (2021) Histone tail conformations: a fuzzy affair with DNA. *Trends Biochem. Sci*, **46**, 564–578.
77. Ai,H., Sun,M., Liu,A., Sun,Z., Liu,T., Cao,L., Liang,L., Qu,Q., Li,Z., Deng,Z. *et al.* (2022) H2B Lys34 ubiquitination induces nucleosome distortion to stimulate Dot1L activity. *Nat. Chem. Biol.*, **18**, 972–980.

78. Kono,H., Shirayama,K., Arimura,Y., Tachiwana,H. and Kurumizaka,H. (2015) Two arginine residues suppress the flexibility of nucleosomal DNA in the canonical nucleosome core. *PLoS One*, **10**, e0120635.
79. Ichikawa,Y., Connelly,C.F., Appleboim,A., Miller,T.C.R., Jacobi,H., Abshiru,N.A., Chou,H.-J., Chen,Y., Sharma,U., Zheng,Y. *et al.* (2017) A synthetic biology approach to probing nucleosome symmetry. *Elife*, **6**, e28836.
80. Dacher,M., Tachiwana,H., Horikoshi,N., Kujirai,T., Taguchi,H., Kimura,H. and Kurumizaka,H. (2019) Incorporation and influence of Leishmania histone H3 in chromatin. *Nucleic Acids Res.*, **47**, 11637–11648.
81. Attar,N., Campos,O.A., Vogelauer,M., Cheng,C., Xue,Y., Schmollinger,S., Salwinski,L., Mallipeddi,N.V., Boone,B.A., Yen,L. *et al.* (2020) The histone H3-H4 tetramer is a copper reductase enzyme. *Science*, **369**, 59–64.
82. Liu,Y., Bisio,H., Toner,C.M., Jeudy,S., Philippe,N., Zhou,K., Bowerman,S., White,A., Edwards,G., Abergel,C. *et al.* (2021) Virus-encoded histone doublets are essential and form nucleosome-like structures. *Cell*, **184**, 4237–4250.
83. Böhm,V., Hieb,A.R., Andrews,A.J., Gansen,A., Rocker,A., Tóth,K., Luger,K. and Langowski,J. (2011) Nucleosome accessibility governed by the dimer/tetramer interface. *Nucleic Acids Res.*, **39**, 3093–3102.
84. Taguchi,H., Horikoshi,N., Arimura,Y. and Kurumizaka,H. (2014) A method for evaluating nucleosome stability with a protein-binding fluorescent dye. *Methods*, **70**, 119–126.
85. Lee,K.M. and Hayes,J.J. (1997) The N-terminal tail of histone H2A binds to two distinct sites within the nucleosome core. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 8959–8964.
86. Obado,S.O., Bot,C., Nilsson,D., Andersson,B. and Kelly,J.M. (2007) Repetitive DNA is associated with centromeric domains in Trypanosoma brucei but not Trypanosoma cruzi. *Genome Biol.*, **8**, R37.
87. Wickstead,B., Ersfeld,K. and Gull,K. (2004) The small chromosomes of Trypanosoma brucei involved in antigenic variation are constructed around repetitive palindromes. *Genome Res.*, **14**, 1014–1024.
88. Kalashnikova,A.A., Porter-Goff,M.E., Muthurajan,U.M., Luger,K. and Hansen,J.C. (2013) The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J.R. Soc. Interface.*, **10**, 20121022.
89. Spangler,C.J., Skrajna,A., Foley,C.A., Nguyen,A., Budziszewski,G.R., Azzam,D.N., Arteaga,E.C., Simmons,H.C., Smith,C.B., Wesley,N.A. *et al.* (2023) Structural basis of paralog-specific KDM2A/B nucleosome recognition. *Nat. Chem. Biol.*, **19**, 624–632.
90. Sato,S., Takizawa,Y., Hoshikawa,F., Dacher,M., Tanaka,H., Tachiwana,H., Kujirai,T., Iikura,Y., Ho,C.-H., Adachi,N. *et al.* (2021) Cryo-EM structure of the nucleosome core particle containing *Giardia lamblia*histones. *Nucleic Acids Res.*, **49**, 8934–8946.
91. Mattiroli,F., Bhattacharyya,S., Dyer,P.N., White,A.E., Sandman,K., Burkhart,B.W., Byrne,K.R., Lee,T., Ahn,N.G., Santangelo,T.J. *et al.* (2017) Structure of histone-based chromatin in Archaea. *Science*, **357**, 609–612.
92. Valencia-Sánchez,M.I., Abini-Agbomson,S., Wang,M., Lee,R., Vasilyev,N., Zhang,J., De Ioannes,P., La Scola,B., Talbert,P., Henikoff,S. *et al.* (2021) The structure of a virus-encoded nucleosome. *Nat. Struct. Mol. Biol.*, **28**, 413–417.
93. Hecker,H., Bender,K., Betschart,B. and Modespacher,U.-P. (1989) Instability of the nuclear chromatin of procyclic Trypanosoma brucei brucei. *Mol. Biochem. Parasitol.*, **37**, 225–234.
94. Hecker,H., Betschart,B., Bender,K., Burri,M. and Schlimme,W. (1994) The chromatin of trypanosomes. *Int. J. Parasitol.*, **24**, 809–819.
95. Ogbadoyi,E., Ersfeld,K., Robinson,D., Sherwin,T. and Gull,K. (2000) Architecture of the Trypanosoma brucei nucleus during interphase and mitosis. *Chromosoma*, **108**, 501–513.
96. Wedel,C., Forstner,K.U., Derr,R. and Siegel,T.N. (2017) GT-rich promoters can drive RNA pol II transcription and deposition of H2A.Z in African trypanosomes. *EMBO J.*, **36**, 2581–2594.
97. Maree,J.P., Povelones,M.L., Clark,D.J., Rudenko,G. and Patterton,H.G. (2017) Well-positioned nucleosomes punctuate polycistronic pol II transcription units and flank silent VSG gene arrays in Trypanosoma brucei. *Epigenet. Chromatin*, **10**, 14.
98. Takizawa,Y., Ho,C.-H., Tachiwana,H., Matsunami,H., Kobayashi,W., Suzuki,M., Arimura,Y., Hori,T., Fukagawa,T., Ohi,M.D. *et al.* (2020) Cryo-EM structures of centromeric tri-nucleosomes containing a central CENP-A nucleosome. *Structure*, **28**, 44–53.
99. Akiyoshi,B. and Gull,K. (2014) Discovery of unconventional kinetochores in kinetoplastids. *Cell*, **156**, 1247–1258.
100. Bender,K., Betschart,B., Marion,C., Michalon,P. and Hecker,H. (1992) Structural differences between the chromatin of procyclic Trypanosoma brucei brucei and of higher eukaryotes as probed by immobilized trypsin. *Acta Trop.*, **52**, 69–78.
101. Kireeva,M.L., Hancock,B., Cremona,G.H., Walter,W., Studitsky,V.M. and Kashlev,M. (2005) Nature of the nucleosomal barrier to RNA polymerase II. *Mol. Cell*, **18**, 97–108.
102. Farnung,L., Vos,S.M. and Cramer,P. (2018) Structure of transcribing RNA polymerase II-nucleosome complex. *Nat. Commun.*, **9**, 5432.
103. Kujirai,T., Ehara,H., Fujino,Y., Shirouzu,M., Sekine,S.-i. and Kurumizaka,H. (2018) Structural basis of the nucleosome transition during RNA polymerase II passage. *Science*, **362**, 595–598.
104. Farnung,L., Ochmann,M., Garg,G., Vos,S.M. and Cramer,P. (2022) Structure of a backtracked hexasomal intermediate of nucleosome transcription. *Mol. Cell*, **82**, 3126–3134.
105. Martinez-Calvillo,S., Romero-Meza,G., Vizuet-de-Rueda,J.C., Florencio-Martinez,L.E., Manning-Cela,R. and Nepomuceno-Mejia,T. (2018) Epigenetic regulation of transcription in trypanosomatid protozoa. *Curr. Genomics*, **19**, 140–149.
106. de Lima,L.P., Poubel,S.B., Yuan,Z.-F., Rosón,J.N., Vitorino,F.N.d.L., Holetz,F.B., Garcia,B.A. and da Cunha,J.P.C. (2020) Improvements on the quantitative analysis of Trypanosoma cruzi histone post translational modifications: study of changes in epigenetic marks through the parasite's metacyclogenesis and life cycle. *J. Proteomics*, **225**, 103847.
107. de Almeida,R.F., Fernandes,M. and de Godoy,L.M.F. (2021) An updated map of Trypanosoma cruzi histone post-translational modifications. *Sci Data*, **8**, 93.
108. Anderson,C.J., Baird,M.R., Hsu,A., Barbour,E.H., Koyama,Y., Borgnia,M.J. and McGinty,R.K. (2019) Structural basis for recognition of ubiquitylated nucleosome by Dot1L methyltransferase. *Cell Rep.*, **26**, 1681–1690.
109. Valencia-Sánchez,M.I., De Ioannes,P., Wang,M., Vasilyev,N., Chen,R., Nudler,E., Armache,J.-P. and Armache,K.-J. (2019) Structural basis of Dot1L stimulation by histone H2B lysine 120 ubiquitination. *Mol. Cell*, **74**, 1010–1019.
110. Worden,E.J., Hoffmann,N.A., Hicks,C.W. and Wolberger,C. (2019) Mechanism of cross-talk between H2B ubiquitination and H3 methylation by Dot1L. *Cell*, **176**, 1490–1501.
111. Davey,G.E., Adhireksan,Z., Ma,Z., Riedel,T., Sharma,D., Padavattan,S., Rhodes,D., Ludwig,A., Sandin,S., Murray,B.S. *et al.* (2017) Nucleosome acidic patch-targeting binuclear ruthenium compounds induce aberrant chromatin condensation. *Nat. Commun.*, **8**, 1575.
112. Cabral,W.F., Machado,A.H.L. and Santos,G.M. (2016) Exogenous nucleosome-binding molecules: a potential new class of therapeutic drugs. *Drug Discov. Today*, **21**, 707–711.
113. Kasozi,K.I., MacLeod,E.T., Ntulume,I. and Welburn,S.C. (2022) An update on African trypanocide pharmaceutics and resistance. *Front. Vet. Sci.*, **9**, 828111.
114. Dodonova,S.O., Zhu,F., Dienemann,C., Taipale,J. and Cramer,P. (2020) Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function. *Nature*, **580**, 669–672.