

ARTICLE

<https://doi.org/10.1038/s41467-019-12007-w>

OPEN

Retroviral integration into nucleosomes through DNA looping and sliding along the histone octamer

Marcus D. Wilson ^{1,7,10}, Ludovic Renault^{1,8,10}, Daniel P. Maskell ^{2,9}, Mohamed Ghoneim^{3,4}, Valerie E. Pye ², Andrea Nans⁵, David S. Rueda ^{3,4}, Peter Cherepanov^{2,6} & Alessandro Costa ¹

Retroviral integrase can efficiently utilise nucleosomes for insertion of the reverse-transcribed viral DNA. In face of the structural constraints imposed by the nucleosomal structure, integrase gains access to the scissile phosphodiester bonds by lifting DNA off the histone octamer at the site of integration. To clarify the mechanism of DNA looping by integrase, we determined a 3.9 Å resolution structure of the prototype foamy virus intasome engaged with a nucleosome core particle. The structural data along with complementary single-molecule Förster resonance energy transfer measurements reveal twisting and sliding of the nucleosomal DNA arm proximal to the integration site. Sliding the nucleosomal DNA by approximately two base pairs along the histone octamer accommodates the necessary DNA lifting from the histone H2A-H2B subunits to allow engagement with the intasome. Thus, retroviral integration into nucleosomes involves the looping-and-sliding mechanism for nucleosomal DNA repositioning, bearing unexpected similarities to chromatin remodelers.

¹Macromolecular Machines Laboratory, The Francis Crick Institute, NW1 1AT London, UK. ²Chromatin structure and mobile DNA Laboratory, The Francis Crick Institute, London NW1 1AT, UK. ³Single Molecule Imaging Group, MRC London Institute for Medical Science, London W12 0NN, UK. ⁴Molecular Virology, Department of Medicine, Imperial College London, London W12 0NN, UK. ⁵Structural Biology Science Technology Platform, The Francis Crick Institute, London NW1 1AT, UK. ⁶Department of Medicine, Imperial College London, St-Mary's Campus, Norfolk Place, London W2 1PG, UK. ⁷Present address: Wellcome Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, UK. ⁸Present address: NeCEN, University of Leiden, 2333CC Leiden, Netherlands. ⁹Present address: Faculty of Biological Sciences, Leeds LS2 9JT, UK. ¹⁰These authors contributed equally: Marcus D. Wilson, Ludovic Renault. Correspondence and requests for materials should be addressed to D.S.R. (email: david.rueda@imperial.ac.uk) or to P.C. (email: peter.cherepanov@crick.ac.uk) or to A.C. (email: alessandro.costa@crick.ac.uk)

Integration of the reverse-transcribed retroviral genome into a host-cell chromosome is catalysed by integrase (IN), an essential viral enzyme (reviewed in¹). To carry out its function, a multimer of IN assembles on viral DNA (vDNA) ends forming a highly stable nucleoprotein complex, known as the intasome^{2–4}. In its first catalytic step, IN resects 3' ends of the vDNA downstream of the invariant CA dinucleotides (3'-processing reaction). It then utilises the freshly released 3'-hydroxyl groups as nucleophiles to attack a pair of phosphodiester bonds on opposing strands of chromosomal DNA, cleaving host DNA and simultaneously joining it to 3' vDNA ends (strand transfer reaction)^{5,6}.

Many important questions pertaining to the nature of the host-virus transactions on chromatin remain unanswered. In particular, it is unclear what role chromatin structure plays in the integration process. Strikingly, although only a fraction of the nucleosomal DNA surface is exposed within the nucleosome core particle (NCP)^{7–9}, nucleosomal DNA packing does not impede and rather stimulates integration^{10–15}. Because retroviral INs have long been known to prefer bent or distorted targets, bending of DNA as it wraps around the histone octamer was thought to facilitate integration into NCPs^{12,13}. However, recent structural data revealed that retroviral intasomes require target DNA to adopt a considerably sharper deformation than the smooth bend observed on NCPs^{15–19}.

Intasome structures from several retroviral genera have been determined by X-ray crystallography and cryo-EM^{4,17–20}. Despite considerable variability, all intasomes were found to contain the structurally conserved intasome core assembly minimally comprising four IN subunits synapsing a pair of vDNA ends. Depending on the retroviral species, the core assembly can be decorated by a number of additional IN subunits. The nucleoprotein complex from the prototype foamy virus (PFV) contains only a tetramer of IN, making this well-characterised intasome an ideal model to study the basic mechanisms involved in retroviral integration. Recently, we reported a cryo-EM structure of the pre-catalytic PFV intasome engaged with an NCP at 7.8 Å resolution¹⁵. Despite the modest level of detail, the cryo-EM data revealed that intasome induces the sharp bending of the nucleosomal DNA by lifting it off the face of the histone octamer at the site of integration. In doing so, the intasome makes supporting interactions with the H2A-H2B heterodimer and the second gyre of the nucleosomal DNA¹⁵. Due to the limited resolution of the original structure, it was impossible to visualise the conformational rearrangements in the nucleosomal DNA that lead to its disengagement from the nucleosomal core at the site of integration. Thus, it remains to be established whether nucleosomal DNA deformation at the integration site is merely accommodated by local deformation of the duplex DNA structure, or it rather involves global repositioning of the nucleosomal DNA along the histone octamer. In addition, a systematic analysis is needed to understand potential role of histone tails in intasome engagement.

Herein, we employ a combination of cryo-EM and single-molecule Förster resonance energy transfer (FRET) assay to understand what impact retroviral integration has on the structure of the target NCP. We find that strand transfer causes both nucleosomal DNA looping, as well as sliding by two base pairs along the histone octamer. With our findings we uncover unexpected similarities between the mechanisms of retroviral integration and ATP-dependent chromatin remodelling^{21–23}.

Results

Structure of Intasome-NCP strand-transfer complex. To understand intasome strand transfer into NCPs, we assembled the

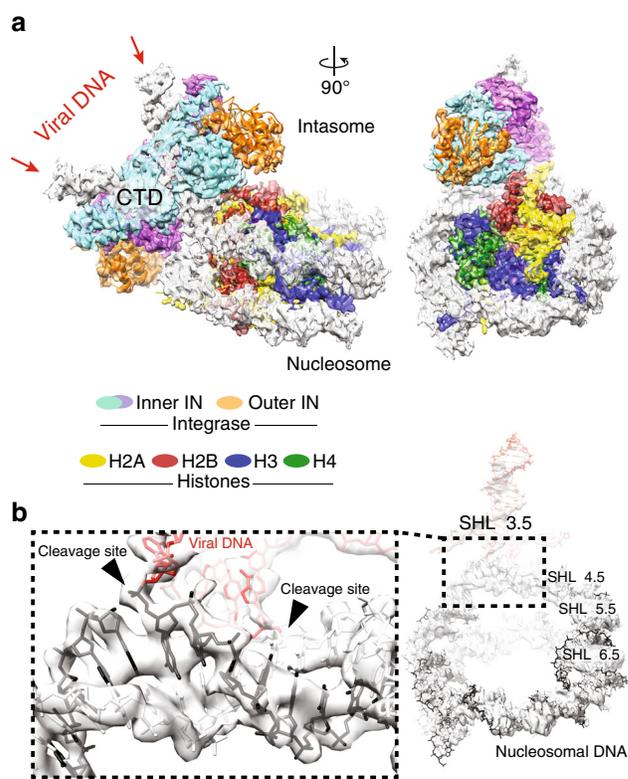


Fig. 1 Intasome-NCP strand-transfer complex visualised by cryo-EM. **a** 3.9-Å resolution structure of the post-catalytic intasome-NCP complex. IN means intasome, CTD means C-terminal domain. **b** Covalently linked viral (red) and nucleosomal (black) DNA. Integration occurred at the superhelical location (SHL) location 3.5

complex of the PFV intasome and the NCP containing a native human DNA sequence (termed D02), previously selected for its ability to form a stable PFV-NCP complex¹⁵. Following isolation by size exclusion chromatography, the intasome-NCP complex was incubated in the presence of Mg^{2+} to facilitate strand transfer¹⁵. We then used cryo-EM imaging and single-particle approaches to determine the structure of the resulting post-catalytic assembly to 3.9 Å resolution (Supplementary Fig. 1, Table 1). Docking known crystallographic coordinates into the cryo-EM map, manual adjustment, and real-space refinement allowed us to generate an atomic model of the Intasome-NCP strand transfer complex.

As previously observed, intasome engages the strongly preferred site on the nucleosomal DNA, at SHL 3.5^{15,24} (Fig. 1). The new structure is overall similar to the original lower-resolution intasome-NCP complex, which was captured in the pre-catalytic state (Fig. 1a), confirming that strand transfer is not accompanied by large conformational rearrangements⁶. According to the atomic model, at the integration site, DNA is lifted by 7 Å from the histone octamer and bent to allow access to the IN catalytic centre, in excellent agreement with the earlier observations based on the crystal structures of the PFV strand transfer complex^{6,16,25} and the lower-resolution intasome-NCP cryo-EM data¹⁵. Local resolution ranges between ~3.5 Å throughout the histone octamer core, and ~4–4.5 Å for nucleosomal DNA, similar to other NCP structures determined by cryo-EM (Supplementary Fig. 1)^{26–28}. Nevertheless, we could confidently model the DNA phosphate backbone for the entire assembly. The integration site on the nucleosomal DNA is sandwiched between the histones and the intasome, resulting in a higher local resolution (~3.7 Å). Notably, a discontinuity in the cryo-EM

Table 1 Data collection and processing information

Parameter	Intasome-NCP	NCP-D02-strep	601 nucleosome
Data Collection			
Microscope	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios
Detector	FEI Falcon II	FEI Falcon III	FEI Falcon III
Acceleration voltage (kV)	300	300	300
Number of micrographs	4916	4182	1300
Frames per micrographs	7	30	30
Frame rate (/s)	4.3	60	60
Dose per frame (e-/pixel)	9.86	1.12	1.24
Accumulated dose (e-/Å ²)	56	28.3	31.3
defocus range (µm)	1.5-3.5	1.5-3.5	1.5-3.5
Frames			
Alignment software	MotionCorr	MotionCor2	MotionCor2
Frames used in final reconstruction	1-7	1-30	2-30
Dose weighting	No	yes	yes
CTF			
Fitting software	CTFFIND3	Gctf	Gctf
Correction	full	full	full
Particles			
Picking software	Xmipp & Relion 1.3	Relion 2.1	Relion 2.1
Picked	989177	1131653	205680
Used in final reconstruction	177155	62196	123123
Alignment			
Alignment software	Relion 1.3	Relion 2.1	Relion 2.1
Initial reference map	EMD-2992	CryoSPARC <i>ab initio</i>	CryoSPARC <i>ab initio</i>
low pass filter limit (Å)	50	50	50
number of iterations	25	25	25
local frame drift correction	yes	no	no
Reconstruction			
Reconstruction software	Relion 1.3	Relion 2.1	Relion 2.1
Box Size	240 × 240 × 240	256 × 256 × 256	256 × 256 × 256
Voxel size (Å)	1.11	1.09	1.09
Symmetry	C1	C1	C2
Resolution limit (Å)	2.22	2.18 Å	2.18 Å
Resolution estimate (Å)	3.9	4.2	3.5
Masking	Yes	Yes	Yes
Sharpening (Å ²)	Bfactor: -146	Bfactor: -150	Bfactor: -110
EMDB ID	EMD-4960	EMD-4692	EMD-4693
Model building			
Number of protein residues	1742	747	
Number of DNA residues	358	284	
Bond length outliers	0.00%	0.00%	
Bond angle outliers	0.02%	0.00%	
Bonds (R.M.S.D)	0.010	0.008	
Angles (R.M.S.D)	1.183	0.856	
Ramachandran favoured/outlier	94.3%/0.00%	96.85%/0%	
Rotamer favoured/outlier	98.5%/0%	99.51/0%	
Clashscore	10.55	4.91	
Model vs Data CC (mask)	0.71	0.85	
Molprobrity score	1.91	1.45	
PDB ID	PDB: 6RNY	PDB: 6ROC	

density resulting from the nucleosomal DNA cleavage at the site of integration (Fig. 1b) confirms that strand transfer has indeed occurred in our nucleoprotein assembly as observed biochemically¹⁵ (Supplementary Fig. 2).

Intasome engages nucleosomal DNA non-symmetrically at two distinct sites: at the strand transfer site, as well as at the opposing gyre, which nestles in the cleft between one catalytic and one outer IN subunit (Fig. 1a). Near the integration site, the C-terminal alpha-helix of histone H2B makes direct contact with the C-terminal domain of one catalytically competent IN subunit, providing corroborating evidence for the previously reported role of PFV IN residues Pro135, Pro239 and Thr240 in engaging the C-terminus of H2B¹⁵. Furthermore, the higher quality of the new cryo-EM map allowed us to build a backbone model for a segment of the N-terminal H2A tail, revealing intimate contacts

of H2A Lys-9 and Arg-11 with the IN C-terminal domain (Fig. 2a). Concordantly, truncation of the first 12, but not 8 H2A residues lead to a reduction of intasome-NCP complex formation (Fig. 2b). Furthermore, Ala substitutions of either H2A at Lys-9 or Arg-11 affect complex stability, while a combination of the two substitutions fully abrogated stable complex formation under conditions of the pull-down assay (Fig. 2c).

Asymmetric reconstruction of the human D02 NCP. Similar to the pre-catalytic complex, our new structure of an intasome-NCP strand-transfer complex features a nucleosomal DNA loop bulging away from the protein octamer by ~7 Å at the integration site. Although occurring at a different superhelical location, the DNA looping is reminiscent of structures of NCPs engaged by

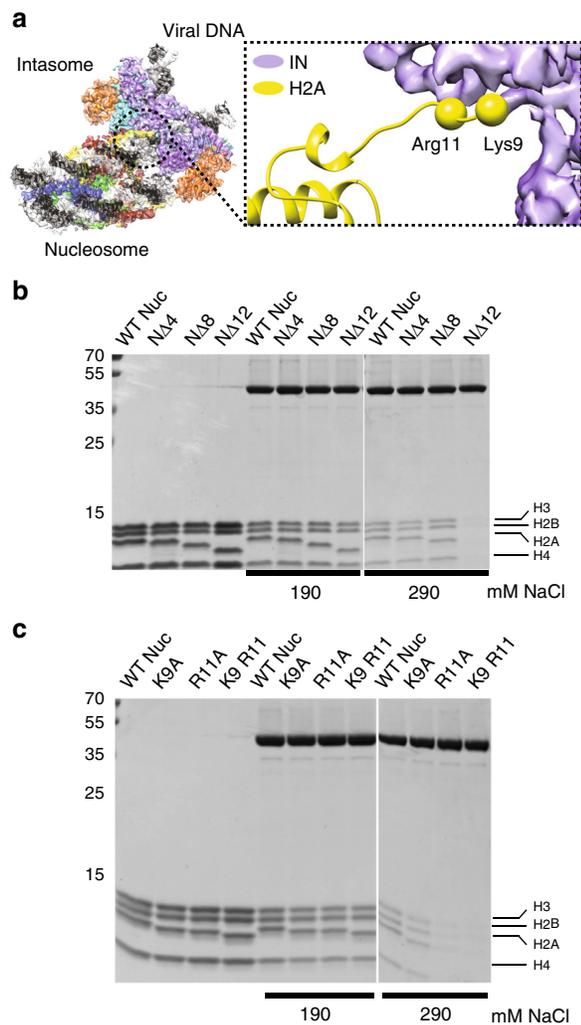


Fig. 2 Intasome interaction with the N-terminal histone H2A tail. **a** Histone H2A residues Lys-9 and Arg-11 play a role in the intasome–NCP interaction. **b, c** Pull-down with immobilised intasome binding to NCP bearing deletions **b** or amino acid substitutions **c** in the H2A N-terminal tail. N Δ 4, N Δ 8, N Δ 12, NCP variants lacking the first 4, 8 or 12 residues for histone H2A

chromatin remodelers such as SWR1. Interestingly, DNA looping by SWR1 is accompanied by both sliding of nucleosomal DNA, as well as histone octamer distortion²². We wanted to test whether intasome-induced looping is compensated by nucleosomal DNA sliding along the histone octamer, as observed for chromatin remodelers. To this end, we decided to directly compare the cryo-EM structure of the intasome–NCP strand-transfer complex with that of an isolated NCP, containing the same native human D02 nucleosomal DNA sequence¹⁵.

Reconstructing a D02 NCP presented a number of significant challenges. Firstly, the NCP containing D02 DNA is less stable than NCPs wrapped with strongly positioning sequences such as Widom 601^{15,29}. Our EM analysis of the isolated NCP D02 revealed that, unlike the intasome complex, D02 NCPs had the tendency to become unravelled, especially in the presence of higher salt measured by the lack of NCP particles in holey grids. However, exposure to mild crosslinking conditions (0.05% glutaraldehyde, 5 min, 4 °C) yielded tractable particles that were visible on open-hole cryo grids. Importantly, mild NCP-crosslinking did not prevent intasome activity as measured in strand-transfer assays (Supplementary Fig. 2). A second challenge was presented by the asymmetry of the D02 DNA sequence,

which leads to the strongly preferred intasome capture at one side of the NCP¹⁵. Thus, to describe any intasome-dependent sliding along the histone octamer, we first had to reconstruct the D02 NCP avoiding two-fold averaging. However, both the histone octamer and the DNA backbone contain a prominent two-fold symmetric character, which strongly influence particle alignment and prevent asymmetric reconstruction. To facilitate asymmetric particle alignment, we introduced a biotin moiety on the end of the DNA arm distal from the integration site and decorated NCPs with streptavidin (Fig. 3a). Critically, streptavidin attachment did not affect NCP stability, nor the ability of intasome to integrate into NCPs (Supplementary Fig. 2). Crosslinked D02 NCPs, imaged by cryo-EM and analysed by two-dimensional (2D) averaging, revealed multiple views of the coin-shaped NCP assemblies (Fig. 3b). Particles appeared decorated by diffuse density projecting from one DNA arm, which we assigned to streptavidin. Free streptavidin particles (~75 kDa) could also be identified amongst the 2D class averages (Supplementary Fig. 3). Next, we used single-particle reconstruction to determine the 4.2-Å resolution structure of NCP–D02–streptavidin complex (Supplementary Fig. 3 and 4). As the streptavidin is linked to the 5' end of a distal DNA arm, it is less ordered than the rest of the assembly, and appears not to be engaged in interactions with the NCP core (Fig. 3c, Supplementary Fig. 3). Therefore, streptavidin helps align particles asymmetrically while seemingly not interfering with the NCP structure.

Originally selected from a genome-wide screen for strong intasome interactors, the D02 DNA sequence allowed isolation of a mono-disperse intasome–NCP complex¹⁵. Detailed inspection of the isolated D02 NCP cryo-EM maps provides insight into intasome selectivity. Firstly, nucleosomal DNA arms appear to be flexible (as detected by inspection of the local resolution map reported in Supplementary Fig. 3, and given the significant number of unwrapped NCPs averaged during analysis). We asked whether the same flexibility could be observed for a NCP containing a strong positioning sequence such as Widom 601, which while serving a good target for strand transfer did not allow formation of long-lived pre-catalytic intasome–NCP complex¹⁵. To this end, we solved a 3.5-Å resolution cryo-EM structure of a Widom 601-wrapped nucleosome containing strongly positioned Widom 601 sequence with 13-bp long linker DNA arms (Supplementary Fig. 5). Only linker DNA fragments display a degree of flexibility in the Widom 601 structure. We postulated at this stage that the flexibility of the D02 NCP DNA arms might facilitate DNA looping, prompting us to further investigate the mechanism.

Another notable feature of the D02 NCP is the limited interaction between DNA and the N-terminal tail of H2A, reflected by poorly defined density contacting nucleosomal DNA at SHL 4.5. This differs for example from our structure of Widom 601 NCP, which shows discrete ordering of H2A N-terminal tail in the minor groove of nucleosomal DNA at the equivalent position, in agreement with previous crystallography and cryo-EM studies^{7,30–32}. We speculate that a loose DNA engagement renders the histone H2A tail available for intasome binding as observed in our strand-transfer complex, hence improving substrate selection (Fig. 3d).

Retroviral integration shifts nucleosomal DNA register. To understand the impact of retroviral integration on NCP architecture, we analysed the structural changes in the NCP that accompany productive engagement with the intasome. Comparison of the intasome–D02 NCP structures prior to and after strand transfer shows that histones undergo relatively minor distortions clustered around the histone H3–H4 dimer on the nucleosomal face proximal to the integration site (Supplementary

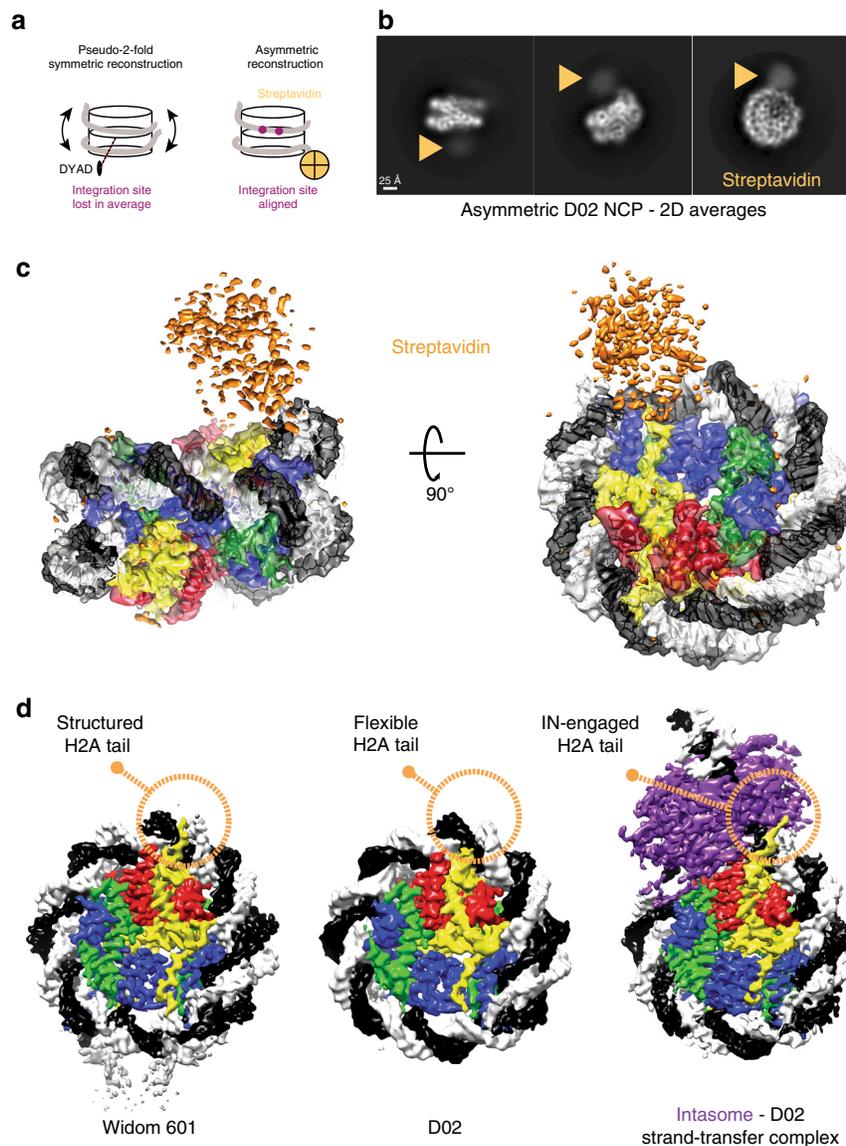


Fig. 3 Asymmetric reconstruction of the isolated D02 NCP. **a** Streptavidin labelling allows asymmetric reconstruction of the D02 NCP, which avoids pseudo-two-fold symmetry averaging of the integration-site DNA. **b** 2D averages of labelled D02 NCP particles reveal a discernible, diffuse density for streptavidin at the nucleosomal DNA arm distal from the integration site. **c** 3D reconstruction of D02 NCP reveals an asymmetric streptavidin label decorating one of two nucleosomal DNA arms. **d** Unlike for Widom 601, the D02 NCP contains limited density for the N-terminal H2A tail, indicating that this element is flexible and available for intasome engagement. In fact, in the Intasome-NCP strand-transfer complex the N-terminal H2A tail interacts with IN, stabilising the interaction (also see Fig. 2b, c)

Fig. 6). Conversely, in our atomic model DNA looping at the integration site is compensated by a significant change in nucleosomal DNA register, with the nucleosomal DNA arm proximal to the integration site shifting by 2 bp (Fig. 4a). This shift in register extends from SHL 7 to SHL 2.5, where an interaction with H3 element L1 appears to hold DNA in place and limit downstream sliding of the double helix (Fig. 4b, Supplementary Movie 1).

To validate the DNA-register change observed in our structural models, we turned to a single-molecule FRET assay. We used a Cy3 donor to label the 5'-terminal end of the nucleosomal DNA closest to the integration site, and a Cy5-maleimide-cysteine acceptor engineered at position 119 of H2A (Fig. 5a). Histone labelling was optimised to yield approximately one fluorophore per octamer. Surface-immobilised NCPs were imaged by FRET in the absence or presence of the intasome and/or Mg^{2+} (Fig. 5b). In reconstituted NCPs, single H2A labels were found either

proximal to, or distal from, the Cy3-modified DNA end. The main energy transfer group deriving from the proximal fluorophore pair centred around 0.95 FRET efficiency, while the second distal fluorophore pair peak centred around 0.37 transfer efficiency (Supplementary Fig. 7A). We focused our analysis on the 0.95 FRET group, as any shift in nucleosomal DNA register would cause more pronounced changes in FRET efficiency in this population. In all tested conditions, FRET efficiency was stable, with a minor population (~10%) of traces exhibiting slight changes in FRET intensity (Fig. 5c, Supplementary Fig. 7B). Supplementing the NCP with intasome or Mg^{2+} did not result in any significant FRET change (Fig. 5d, e). However, when strand transfer was induced by adding both intasome and Mg^{2+} , a separate, ~0.8 FRET population appeared (Fig. 5f). This second population is consistent with a shift in register of the DNA moving away from the K119C-Cy5 H2A residue (Fig. 5a). These results are in good agreement with our comparative cryo-EM

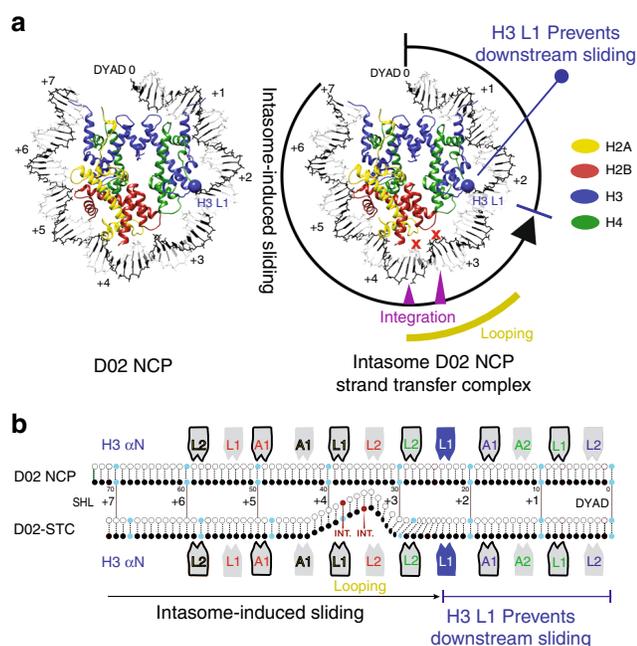


Fig. 4 Integration-promoted DNA sliding observed by cryo-EM. **a** On the left, slabbed view of the isolated D02 NCP. Histone H3 L1 element is highlighted with a blue ball. On the right, DNA looping required for retroviral integration causes a shift in the DNA register, which extends from SHL 7 to 2. Histone H3 L1 element prevents downstream DNA sliding. **b** Schematic representation of integration-induced NCP remodelling

analyses indicating that intasome-mediated looping required for integration promotes sliding of nucleosomal DNA (Fig. 4a). In fact, the observed drop in FRET efficiency indicates a small but significant shift in the DNA register that corresponds to less than 4 bp, according to a calibration previously obtained with Widom 601 NCPs²². Crystallographic and cryo-EM structures of pre-catalytic assemblies of intasome bound to DNA or nucleosomes established that target capture alone leads to DNA bending and nucleosomal DNA remodelling^{15,16}. The new post-catalytic intasome-NCP structure reported shows no change of DNA looping at the integration site after strand transfer. However, in our single molecule experiments, a drop in FRET efficiency was only observed in the presence of Mg^{2+} -required for catalysis. Although this observation was initially surprising to us, we note that the pre-catalytic intasome-nucleosome complex used in our earlier work¹⁵ was purified under elevated ionic strength conditions to enrich for higher affinity productive interaction³³. Indeed, of the two symmetry-related D02 SHL ± 3.5 intasome binding sites, near equally targeted in a bulk strand transfer assay, only one was occupied in the purified material¹⁵. Conditions of the single molecule FRET used here were more similar to the bulk strand transfer assay, allowing for detection of transient interactions and fixation of productive complexes in the absence and presence of Mg^{2+} , respectively. We speculate that most intasome complexes observed in the absence of the metal ion cofactor result from transient scanning¹⁵ interaction, which alone is unlikely to result in DNA deformation. Alternatively, DNA perturbation and sliding could be functionally distinct steps. This could be mediated by histone buffering of DNA displacement, as observed previously³⁴. In this model, tension in the lifted DNA at the integration site could be partially accommodated by protein-DNA contacts within the target capture complex, without an immediate shift in DNA, which would occur upon full strand transfer catalysed upon addition of Mg^{2+} .

Discussion

Over the last 35 years macromolecular crystallography has provided several high-resolution views of the NCP and its binding partners. These efforts led to describing the NCP architecture at an atomic level^{7–9}, explained how DNA sequence can influence wrapping of the double helix³⁵, and how common docking sites on the histone octamer are recognised by different interactors^{36–40}. Over the last four years, cryo-EM has started to provide a dynamic view of the NCP^{26,41–44}. Recent data indicated that NCPs are more flexible in solution, with the histone octamer visiting more compacted or extended states, compared with a nucleosome trapped in a crystal lattice⁴⁵. NCP unwrapping has been visualised with cryo-EM, for example in the context of the hexasome, which is an NCP with partially unpeeled DNA, due to the loss of one H2A/H2B dimer²⁸. Spectacular views of progressively unwrapped NCPs have been obtained for transcribing RNA polymerase II captured during NCP passage^{46,47}. Moreover, cryo-EM provided the first glimpses of ATP-dependent NCP translocation through a mechanism involving DNA looping and sliding along the histone octamer^{22,41,48–54}.

Our high-resolution view of a post-catalytic intasome-NCP complex provides an example of a local remodelling of nucleosomal DNA. Although previous work established formation of a DNA loop during productive intasome-NCP interaction, it was not clear whether the loop is accommodated by partial underwinding of flanking DNA or through a shift in nucleosomal DNA register. Because IN must catalyse only one strand transfer event and does not need to cycle between states on the chromatin, it does not depend on a power source, unlike ATP-driven translocases and nucleosome remodelers³⁴. Therefore, all conformational DNA rearrangements are offset by energy released with the formation of the intasome-NCP interface. Nevertheless, similarities with the mechanism of DNA translocation of chromatin remodelers can be identified. In fact, in both systems, DNA is looped out of the histone core, causing a compensatory register shift of the double helix wrapped around the octamer. Nucleosomal DNA looping at SHL 3.5 is required for access to the IN active site¹⁵, and causes DNA sliding around the histone octamer, with global repositioning extending from SHL 7 to SHL 2. At this site, histone H3 element L1 holds the sugar-phosphate backbone in place, preventing any further downstream shift in DNA register (Fig. 4b, Supplementary Movie 1). Using cryo-EM, Kurumizaka and colleagues have recently shown that the same H3 L1-DNA interaction stalls RNA polymerase II during nucleosome passage⁴⁶. ATP-powered translocases such as Swr1 and Snf2 have been observed to engage and loop out SHL 2 DNA, disrupting the H3 L1-DNA interaction^{22,48,55}. It is tempting to speculate that the concerted action of intasome and SHL 2 remodelers could act synergistically during DNA unpeeling and strand-transfer complex disassembly, required to complete retroviral integration.

Methods

Intasome purification. The intasome was assembled using recombinant PFV IN and double stranded synthetic oligonucleotides mimicking the pre-processed U5 end of the vDNA as previously described^{4,15}. Briefly, hexahistidine-tagged IN was overexpressed in BL-21 CodonPlus RIL cells (Agilent). Cells were lysed in 25 mM Tris-HCl pH 7.4, 0.5 M NaCl, 1 mM PMSF by sonication; clarified lysate supplemented with 20 mM imidazole was applied to packed, equilibrated Ni-NTA resin (Qiagen). The resin and washed extensively in lysis buffer supplemented with 200 mM imidazole and protein-containing fractions were supplemented with 5 mM DTT. The hexahistidine-tag was cleaved by incubation with human rhinovirus 14 3C protease. The protein, diluted to reduce the NaCl concentration to 200 mM, was loaded onto a HiTrap Heparin column (GE Healthcare). IN was eluted using a linear gradient of 0.25–1 M NaCl. IN-containing fractions were concentrated and further purified by size exclusion chromatography through a Superdex-200 column (GE Healthcare), equilibrated in 25 mM Tris pH 7.4, 0.5 M NaCl. Protein, supplemented with 10% glycerol and 10 mM DTT, was

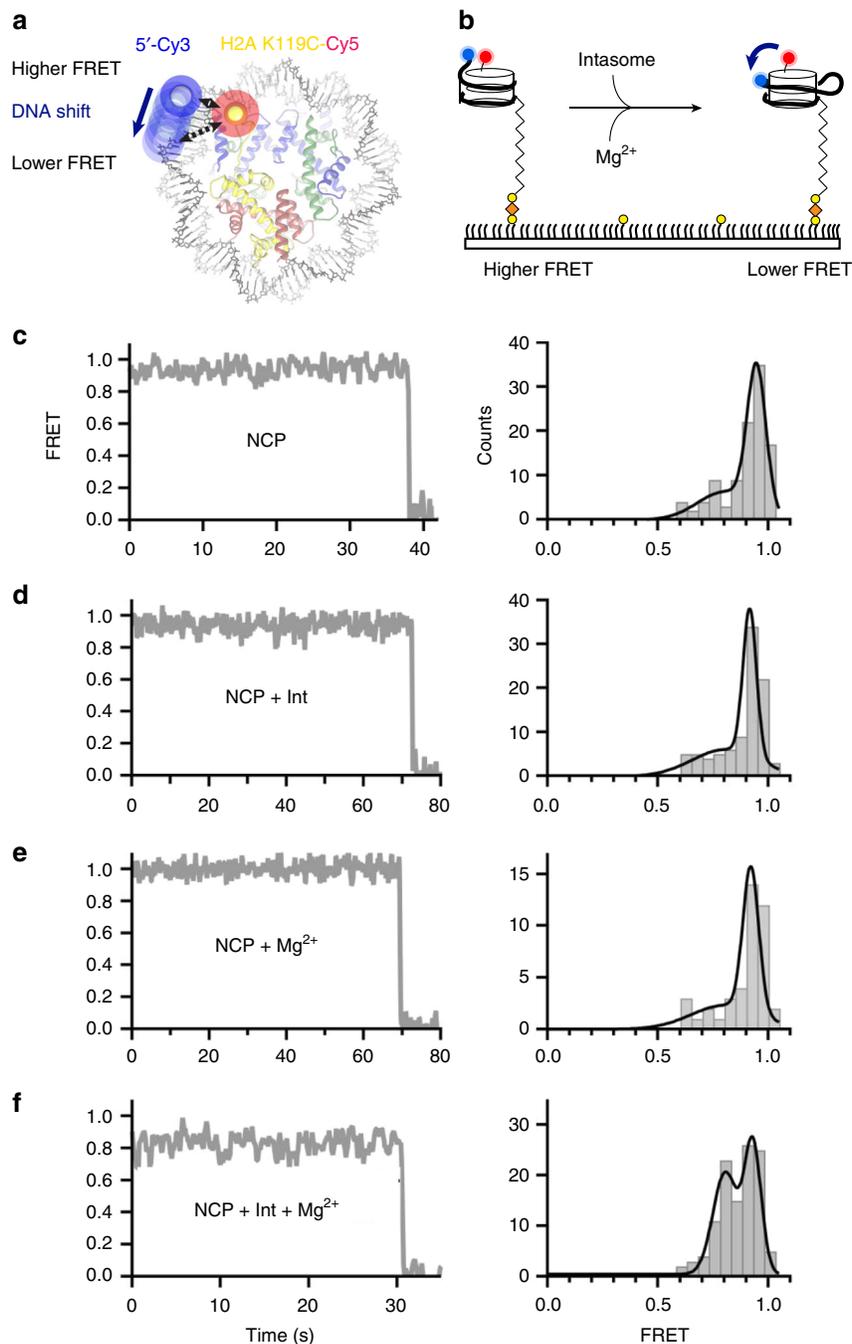


Fig. 5 Integration-promoted nucleosomal DNA sliding observed by single molecule FRET. **a** Fluorescently labelled NCP: labelled octamer (**Cy5** on H2A, red) wrapped with biotinylated labelled DNA (**Cy3** on exit site, blue). **b** NCPs are surface-immobilised on neutravidin (orange) coated biotin-PEG (yellow) slides. Intasome-induced translocation in the presence of magnesium is detected as a FRET decrease. Representative single-molecule FRET trajectories (left) and histograms (right) of proximal-only labelled NCPs **c** without intasome or magnesium ($N = 105$), **d** in the presence of intasome ($N = 93$), **e** in the presence of magnesium ($N = 42$), and **f** in presence of both intasome and magnesium ($N = 115$). A significant population shift to ~ 0.8 FRET in the presence of both intasome and magnesium. Data were collected at 100 ms/frame (10 Hz) and smoothed with a 3-point moving average. Black lines in FRET histogram are fits to two Gaussian distributions

concentrated to 10 mg/ml, as estimated by spectrophotometry at 280 nm and stored at -80°C .

To assemble the intasome a mixture containing 120 μM PFV IN and 20 μM pre-annealed DNA oligonucleotides 5'-TGCGAAATTCATGACA and 5'-ATTGTCATGGAATTCGCA (IDT) in 500 mM NaCl was dialysed against 50 mM BisTris propane-HCl pH 7.45, 200 mM NaCl, 40 μM ZnCl_2 , 2 mM DTT for 16 h at 18°C . A list of all oligos is provided in Supplementary Table 1. Following dialysis, the assembly reaction, supplemented with NaCl to a final concentration of 320 mM, was incubated on ice for 1 h prior to purification on Superdex-200 column in 25 mM Bis-Tris propane-HCl pH 7.45, 320 mM NaCl. Purified intasome, concentrated by ultrafiltration, was kept on ice for immediate use.

NCP formation. NCPs were assembled essentially as described^{15,56}. Briefly Human H2A, H2A K119C, H2B, H3.3, H3.1 C96SC110A and H4 were over-expressed in *E. coli* and purified from inclusion bodies. Histones were refolded from denaturing buffer through dialysis against 10 mM Tris-HCl pH 7.5, 2 M NaCl, 5 mM beta-mercaptoethanol, 1 mM EDTA buffer, and octamers were purified by size exclusion chromatography over a Superdex-200 column (GE Healthcare). DNA fragments for wrapping NCPs (171-bp Widom 601 DNA, 145-bp D02 DNA or D02 DNA appended with biotin and fluorophores) were generated by PCR using Pfu polymerase and HPLC-grade oligonucleotides (IDT). PCR products generated in 96-well plates ($384 \times 100 \mu\text{l}$) were pooled, filtered and purified on a ResorceQ column as described¹⁵. NCPs were assembled by salt dialysis as described^{15,30,56} and heat

repositioned at 37 °C for 30 min. D02 containing NCPs were further purified using a PrepCell apparatus with a 5% polyacrylamide gel (BioRad).

NCP-streptavidin complex formation. Purified *Streptomyces avidinii* streptavidin powder (Sigma-Aldrich) was resuspended in 20 mM HEPES-NaOH pH 7.5, 150 mM NaCl at a final concentration of 35 μ M. A derivative of D02 DNA was used for NCP reconstitution, containing a 5' biotin moiety on the exit arm distal from the intasome-engagement site. To form the NCP-streptavidin complex, biotinylated D02 NCP (0.5 μ M) was incubated with 0.3 μ M streptavidin for 10 minutes at room temperature in 20 mM HEPES pH7.5, 150 mM NaCl, 1 mM DTT, 1 mM EDTA.

EM sample preparation. The intasome-D02 NCP complex was formed and purified by size exclusion chromatography as previously described¹⁵. Briefly, 200 μ g of D02 NCP and 200 μ g of PFV intasome were mixed in 25 mM Bis-Tris Propane, 320 mM NaCl, prior to application on Superdex 200 10/300 column. To allow strand transfer, the complex was incubated in the presence of 5 mM MgCl₂ for 30 min at room temperature. Cryo-EM sample preparation was performed as follows: 4 μ l of the integration reaction were applied to plasma cleaned C-Flat 1/1 400 mesh grids; after 1 min incubation, grids were double side blotted for 3.5 s using a CP3 cryo-plunger (Gatan), operated at 80% humidity, and quickly plunge-frozen into liquid ethane. Ice quality was checked using a JEOL-2100 Lab6 operated at 120 kV, using a 914 side-entry cryo-holder (Gatan), and images were recorded on an UltraScan 4kx4k camera (Gatan). The best cryo-grids were retrieved, stored in liquid nitrogen and later shipped in a dry-shipper to NeCEN (University of Leiden, The Netherlands). At NeCEN, grids were loaded into a Cs corrected Titan Krios microscope and the data was collected over two different sessions using the EPU software (ThermoFisher Scientific). Images were recorded at a nominal magnification of 59,000 X on Falcon II direct electron detector yielding a pixel size of 1.12 Å/pixel with a defocus range of -1.5 to -3.5 μ m. Data were collected as movies of 7 frames over 1.6 s giving a total applied dose of 56 electrons/Å². A total of 4,916 movies were collected.

The D02 NCP biotin-streptavidin complex was gently cross-linked with 0.05% glutaldehyde at room temperature for 5 min, prior to quenching with 50 mM TrisHCl pH 7.5. The complex was concentrated and buffer exchanged using a 50-kDa MWCO spin concentrator (Amicon) into 10 mM Tris-HCl pH 7, 20 mM NaCl, 1 mM EDTA, 1 mM DTT; 3.5 μ l sample at 80 ng/ μ l (DNA concentration based on spectrophotometry) was applied to Quantifoil 2/2 grids, with fresh carbon pre-evaporated onto the grids to better control ice thickness. Grids were glow discharged at 40 mA for 1 min. Sample was blotted in a Vitrobot Mark IV using -1 offset, 15 s wait time and 2.5 s blot at 4 °C and 100% humidity, before plunge-freezing in liquid ethane. Grids were stored in liquid nitrogen prior to loading on a Titan Krios operated at 300 kV. Data was acquired using a Falcon III detector operating in counting mode using a pixel size of 1.09 Å, a total dose of 30 electrons/Å² and a defocus range from -1.5 to -3.5 μ m. A total of 4,182 movies were collected automatically using the EPU software (ThermoFisher Scientific). The Widom 601 NCP sample was applied to freshly glow discharged Quantifoil 2/2 grids and sample was blotted in a Vitrobot Mark IV using -1 offset, 10 s wait time, 3.5 s blot at 4 °C and 100% humidity, before plunge-freezing in liquid ethane. Data were acquired using a Falcon III detector operating in counting mode using a pixel size of 1.09 Å and total dose of 30 electrons/Å². A total of 1,300 Micrographs were collecting using automated EPU software.

Cryo-EM image processing. For the intasome-D02 NCP complex dataset (Supplementary Fig. 1), movie frames were corrected for beam-induced drift⁵⁷ and a sum of each aligned movie was used in the first steps of image processing. All movies showing any remaining drift or containing ice were discarded at this stage, and only the best 3,125 movies were selected for further image processing. First, 989,177 particles were automatically picked using Xmipp⁵⁸ and Relion version 1.3⁵⁹. Contrast transfer function parameters were estimated using CTFFIND3⁶⁰, and all 2D and 3D classifications and 3D refinements were performed using RELION⁵⁹. After 2 rounds of 25 iterations of 2D classification, 335,989 particles remained and were subjected to 3D classification using the pre-catalytic intasome-NCP map¹⁵, filtered to 50 Å resolution, as a starting model. To speed up calculations, 8 classes were generated with a 15° angular sampling. The best 3 classes were merged into one 232,000 particles dataset. 3D refinement of this subset yielded a 4.7 Å map. A second round of 3D classification step was performed with 4 classes and a finer 7.5° angular sampling. The best 3 classes were merged together for a total of 177,155 particles. Refinement of this dataset yielded a 4.2 Å map. Statistical movie processing was then performed, as described previously⁶¹ and the resulting map reached 3.9 Å resolution after correction for the modulation transfer function and sharpening⁶². Resolutions are reported according to the “gold-standard” Fourier Shell Correlation, using the 0.143 criterion⁶³.

For the D02-NCP-Streptavidin and Widom 601 NCP datasets (Supplementary Figs. 3-5) all micrographs were motion-corrected using MotionCorr2 using all frames (D02-NCP-Streptavidin) or removing the first frame (Widom 601 NCP). CTF parameters were estimated using Gctf⁶⁴ and poor micrographs were discarded. Particles were picked in RELION-2.1 using reference classes obtained from a manually-picked, 50-micrograph dataset. Two rounds of 2D classification

were performed to discard poorly averaging particles. 3D classification was performed using a 50 Å, low pass filtered initial model, based on results from an *ab initio* reconstruction derived from cryoSPARC⁶⁵. For the Widom 601 NCP, particles contributing to 3D classes with discernible secondary-structure features were pooled and refined using a spherical mask, and postprocessed in RELION-2.1⁶⁶ resulting in a 3.8 Å (C1 symmetry applied) or 3.5 Å resolution (C2 symmetry applied). For the D02-NCP-Streptavidin, a relatively smaller percentage of particles contributed to subnanometre-resolution 3D averages. This is likely because of evident flexibility of the both the exit nucleosomal DNA and the streptavidin group. To help drive streptavidin alignment and avoid artificial NCP symmetrisation, a loose mask was used in a subsequent round of 3D classification, encompassing both NCP and streptavidin. The resulting asymmetric reconstruction yielded a reconstruction with 4.6 Å (no mask) or 4.2 Å resolution (loose mask applied during refinement).

Atomic model docking and refinement. For the NCP-intasome STC complex NCP (3UTB⁶⁷ from PDBredo) and PFV strand transfer complex (3OS0¹⁶) crystal structures were docked in the EM map using Chimera⁶⁸ and clashing DNA segments were removed from the model. In order to refine the voxel/pixel size of the map a series of maps were calculated with voxel/pixel size from 0.9 to 1.15 in steps of 0.01 and the initial model was refined against each map using phenix.real_space_refine⁶⁹ with no additional geometry restraints. The geometry of resulting models was compared, and voxel/pixel size fine-tuned between 1.11 and 1.12 in steps of 0.001. The model refined against the map with voxel/pixel size of 1.111 maintained the best geometry and was used for further model building and refinement. The model was adjusted, and sequence of protein and DNA components matched to the biological sample manually in Coot⁷⁰ and refined using phenix.real_space_refine (Nightly build version 1.10pre-2091)⁷¹ and Namdinator^{72,73}. Additional restraints describing protein secondary structure, DNA base pairing and stacking were used in Phenix. Protein geometry was assessed with Molprobit⁶⁹ and DNA geometry was assessed with 3DNA⁷⁴. For the D02 structure, NCP structure 5MLU was used as the starting model to be independent from the NCP-intasome STC structure. The sequence was adjusted and model manually tweaked in Coot and refined using phenix.real_space_refine (Nightly build version phenix-dev-3374). Fine tuning of the voxel/pixel size was deemed unnecessary as the model refined without issue. Both models have reasonable stereochemistry and are in good agreement with the EM maps.

Single-Molecule FRET experiments. Doubly-labelled nucleosomes were generated with a biotin on distal exit DNA and a single fluorophore donor (Cy3) attached on the proximal exit DNA end, and the acceptor fluorophore (Cy5) at H2A position 119. To generate protein-Cy5-labelled octamers H2A K119C was incorporated into octamers with H3.1 C96SC110A, H2B and H4 as described above, with an additional desalting step in a Zeba Spin column (ThermoFisher, 7 K MWCO) to remove beta-mercaptoethanol. Octamers at 70 μ M (140 μ M of cysteine) were incubated with 5 mM TCEP for 10 min at room temperature. To achieve partial labelling, sulpho-Cy5 maleimide was added at 105 μ M for 1 hour at room temperature. The reaction was quenched by adding 5 mM beta-mercaptoethanol and desalted to remove unreacted dye (ThermoFisher, 7 K MWCO). The extent of labelling was quantified by measuring the 595 nm/280 nm absorbance ratio, as well as by 2D intact mass ESI mass spectrometry, with an estimated labelling efficiency of 68%. D02 DNA was generated by PCR, using oligos containing Biotin-TEG-C18 and Cy3 modifications attached to the 5' termini. The PCR product was purified as described above. Nucleosomes were reconstituted as described above.

Single-Molecule FRET experiments were performed with freshly purified intasome complex. Quartz slides and coverslips were cleaned and passivated with methoxy-PEG-SVA (M_n = 5,000, Laysan Bio, Inc.) containing 10% biotin-PEG-SVA (M_n = 5,000, Laysan Bio, Inc.) in 100 mM sodium bicarbonate, and used to construct a microfluidic channel as described previously⁷⁵. Neutravidin (0.2 mg/ml in 50 mM Tris-HCl, pH 7.5, and 50 mM NaCl) was injected in and incubated for 5 min. Excess neutravidin was washed out with intasome buffer (25 mM bis-Tris propane, pH 7.45, 240 mM NaCl, 4 μ M ZnCl₂ and 1 mM DTT). Biotinylated fluorescently labelled nucleosomes in intasome buffer containing 0.2 mg/ml BSA were surface immobilised by incubation in the microfluidic channel for 5 min. Excess nucleosomes were washed out and immobilised nucleosomes imaged in imaging buffer composed of intasome buffer in addition to 2 mM Trolox, oxygen scavenging system (2.5 mM 3,4-dihydroxybenzoic acid, 250 nM protocatechuate dioxygenase) and 0.2 mg/ml BSA. Experiments were performed in the absence and presence of 500 nM intasome and 5 mM magnesium. Fluorescent molecules were imaged using a custom-built prism-based total-internal reflection fluorescence (TIRF) microscope⁷⁶. All measurements were recorded at room temperature (21 °C) using continuous green laser (532 nm, 2.5 mW) excitation at 100 ms time resolution. Apparent FRET efficiencies were calculated as the ratio of acceptor intensity divided by the sum of acceptor and donor intensities. FRET histograms of labelled nucleosomes were obtained by calculating the mean FRET efficiency of 40–100 trajectories from multiple fields of view, as stated in the figure captions. All experiments have been performed at least twice, on different days and with different combinations of protein preps.

Intasome strand-transfer and pull-down assays. Intasome integration assays were performed as described¹⁵, briefly 5 µg of NCPs were incubated with 1.5 µg of intasome in intasome reaction buffer with and without 5 mM MgCl₂ at 37 °C for 15 min. The reaction was quenched by the addition of 25 mM EDTA and 0.2% SDS, and DNA precipitated after proteinase K digestion. DNA was then separated on 4–12% TBE polyacrylamide gels. Pull-down assays were performed, as previously described¹⁵. Briefly 10 µg of biotinylated intasome was incubated with 10 µg of NCP variants in pull-down buffer with increasing concentrations of sodium chloride. PFV and associated NCPs was immobilised on streptavidin beads (Life technologies), washed extensively and eluted by heating at 37 °C in 1.3× SDS Laemmli buffer. All experiments have been performed at least twice, on different days and with different combinations of protein preps.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Model coordinates for the NCP-D02-streptavidin and Intasome-NCP structures are deposited in the Protein Data Bank under accession code 6RNY and 6ROC respectively. Cryo-EM maps for NCP-D02-streptavidin, NCP-601 and Intasome-NCP are available at the EMDB under codes EMD-4692, EMD-4693 and EMD-4960 respectively. The source data underlying Fig. 2b and c, 5c–f and 7c and Supplementary Figs 2, 3a, 5a, 6a and 7 are provided as a Source Data file. Other data are available from the corresponding authors upon reasonable request.

Received: 14 May 2019 Accepted: 8 August 2019

Published online: 13 September 2019

References

- Lesbats, P., Engelman, A. N. & Cherepanov, P. Retroviral DNA Integration. *Chem. Rev.* **116**, 12730–12757 (2016).
- Li, M., Mizuuchi, M., Burke, T. R. Jr. & Craigie, R. Retroviral DNA integration: reaction pathway and critical intermediates. *EMBO J.* **25**, 1295–1304 (2006).
- Heuer, T. S. & Brown, P. O. Photo-cross-linking studies suggest a model for the architecture of an active human immunodeficiency virus type 1 integrase-DNA complex. *Biochemistry* **37**, 6667–6678 (1998).
- Hare, S., Gupta, S. S., Valkov, E., Engelman, A. & Cherepanov, P. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **464**, 232–236 (2010).
- Gerton, J. L., Herschlag, D. & Brown, P. O. Stereospecificity of reactions catalyzed by HIV-1 integrase. *J. Biol. Chem.* **274**, 33480–33487 (1999).
- Hare, S., Maertens, G. N. & Cherepanov, P. 3'-processing and strand transfer catalysed by retroviral integrase in crystallo. *EMBO J.* **31**, 3020–3028 (2012).
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
- Richmond, T. J. & Davey, C. A. The structure of DNA in the nucleosome core. *Nature* **423**, 145–150 (2003).
- Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D. & Klug, A. Structure of the nucleosome core particle at 7 Å resolution. *Nature* **311**, 532–537 (1984).
- Pryciak, P. M., Sil, A. & Varmus, H. E. Retroviral integration into minichromosomes in vitro. *EMBO J.* **11**, 291–303 (1992).
- Pryciak, P. M. & Varmus, H. E. Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell* **69**, 769–780 (1992).
- Pruss, D., Bushman, F. D. & Wolffe, A. P. Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. *Proc. Natl Acad. Sci. USA* **91**, 5913–5917 (1994).
- Muller, H. P. & Varmus, H. E. DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes. *EMBO J.* **13**, 4704–4714 (1994).
- Wang, G. P., Ciuffi, A., Leipzig, J., Berry, C. C. & Bushman, F. D. HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res.* **17**, 1186–1194 (2007).
- Maskell, D. P. et al. Structural basis for retroviral integration into nucleosomes. *Nature* **523**, 366–369 (2015).
- Maertens, G. N., Hare, S. & Cherepanov, P. The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature* **468**, 326–329 (2010).
- Ballandras-Colas, A. et al. A supramolecular assembly mediates lentiviral DNA integration. *Science* **355**, 93–95 (2017).
- Passos, D. O. et al. Cryo-EM structures and atomic model of the HIV-1 strand transfer complex intasome. *Science* **355**, 89–92 (2017).
- Yin, Z. et al. Crystal structure of the Rous sarcoma virus intasome. *Nature* **530**, 362–366 (2016).
- Ballandras-Colas, A. et al. Cryo-EM reveals a novel octameric integrase structure for betaretroviral intasome function. *Nature* **530**, 358–361 (2016).
- Wigley, D. B. & Bowman, G. D. A glimpse into chromatin remodeling. *Nat. Struct. Mol. Biol.* **24**, 498–500 (2017).
- Willhoft, O. et al. Structure and dynamics of the yeast SWR1-nucleosome complex. *Science* **362**, 7716 (2018).
- Cairns, B. R. Chromatin remodeling: insights and intrigue from single-molecule studies. *Nat. Struct. Mol. Biol.* **14**, 989–996 (2007).
- Mackler, R. M. et al. Nucleosome DNA unwrapping does not affect prototype foamy virus integration efficiency or site selection. *PLoS ONE* **14**, e0212764 (2019).
- Yin, Z., Lapkouski, M., Yang, W. & Craigie, R. Assembly of prototype foamy virus strand transfer complexes on product DNA bypassing catalysis of integration. *Protein Sci.* **21**, 1849–1857 (2012).
- Wilson, M. D. & Costa, A. Cryo-electron microscopy of chromatin biology. *Acta Crystallogr. D. Struct. Biol.* **73**, 541–548 (2017).
- Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**, 1097–1113 (2002).
- Bilokapic, S., Strauss, M. & Halic, M. Histone octamer rearranges to adapt to DNA unwrapping. *Nat. Struct. Mol. Biol.* **25**, 101–108 (2018).
- Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19–42 (1998).
- Wilson, M. D. et al. The structural basis of modified nucleosome recognition by 53BP1. *Nature* **536**, 100–103 (2016).
- Chua, E. Y. et al. 3.9 Å structure of the nucleosome core particle determined by phase-plate cryo-EM. *Nucleic Acids Res.* **44**, 8013–8019 (2016).
- Iwasaki, W. et al. Comprehensive structural analysis of mutant nucleosomes containing lysine to glutamine (KQ) substitutions in the H3 and H4 histone-fold domains. *Biochemistry* **50**, 7822–7832 (2011).
- Jones, N. D. et al. Retroviral intasomes search for a target DNA by 1D diffusion which rarely results in integration. *Nat. Commun.* **7**, 11409 (2016).
- Sabantsev, A., Levendosky, R. F., Zhuang, X., Bowman, G. D. & Deindl, S. Direct observation of coordinated DNA movements on the nucleosome during chromatin remodelling. *Nat. Commun.* **10**, 1720 (2019).
- McGinty, R. K. & Tan, S. Nucleosome structure and function. *Chem. Rev.* **115**, 2255–2273 (2015).
- McGinty, R. K., Henrici, R. C. & Tan, S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**, 591–596 (2014).
- Makde, R. D., England, J. R., Yennawar, H. P. & Tan, S. Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature* **467**, 562–566 (2010).
- Arnaudo, N. et al. The N-terminal acetylation of Sir3 stabilizes its binding to the nucleosome core particle. *Nat. Struct. Mol. Biol.* **20**, 1119–1121 (2013).
- McGinty, R. K. & Tan, S. Recognition of the nucleosome by chromatin factors and enzymes. *Curr. Opin. Struct. Biol.* **37**, 54–61 (2016).
- Zhou, K., Gaullier, G. & Luger, K. Nucleosome structure and dynamics are coming of age. *Nat. Struct. Mol. Biol.* **26**, 3–13 (2018).
- Anderson, C. J. et al. Structural basis for recognition of ubiquitylated nucleosome by Dot1L methyltransferase. *Cell Rep.* **26**, 1681–1690 e5 (2019).
- Worden, E. J., Hoffmann, N. A., Hicks, C. W. & Wolberger, C. Mechanism of cross-talk between H2B ubiquitination and H3 methylation by Dot1L. *Cell* **176**, 1490–1501 e12 (2019).
- Valencia-Sanchez, M. I. et al. Structural basis of Dot1L stimulation by histone H2B Lysine 120 ubiquitination. *Mol. Cell* **74**, 1010–1019 (2019).
- Jang, S. et al. Structural basis of recognition and destabilization of the histone H2B ubiquitinated nucleosome by the DOT1L histone H3 Lys79 methyltransferase. *Genes Dev.* **33**, 620–625 (2019).
- Bilokapic, S., Strauss, M. & Halic, M. Structural rearrangements of the histone octamer translocate DNA. *Nat. Commun.* **9**, 1330 (2018).
- Kujirai, T. et al. Structural basis of the nucleosome transition during RNA polymerase II passage. *Science* **362**, 595–598 (2018).
- Ehara, H. et al. Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science* **363**, 744–747 (2019).
- Liu, X., Li, M., Xia, X., Li, X. & Chen, Z. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* **544**, 440–445 (2017).
- Farnung, L., Vos, S. M., Wigge, C. & Cramer, P. Nucleosome-Chd1 structure and implications for chromatin remodelling. *Nature* **550**, 539–542 (2017).
- Eustermann, S. et al. Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* **556**, 386–390 (2018).
- Ayala, R. et al. Structure and regulation of the human INO80-nucleosome complex. *Nature* **556**, 391–395 (2018).
- Sundaramoorthy, R. et al. Structure of the chromatin remodelling enzyme Chd1 bound to a ubiquitylated nucleosome. *Elife* **7**, e35720 (2018).

53. Sundaramoorthy, R. et al. Structural reorganization of the chromatin remodeling enzyme Chd1 upon engagement with nucleosomes. *Elife* **6**, e22510 (2017).
54. Yan, L., Wu, H., Li, X., Gao, N. & Chen, Z. Structures of the ISWI-nucleosome complex reveal a conserved mechanism of chromatin remodeling. *Nat. Struct. Mol. Biol.* **26**, 258–266 (2019).
55. Li, M. et al. Mechanism of DNA translocation underlying chromatin remodelling by Snf2. *Nature* **567**, 409–413 (2019).
56. Dyer, P. N. et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzym.* **375**, 23–44 (2004).
57. Li, X. et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
58. Sorzano, C. O. et al. XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol.* **148**, 194–204 (2004).
59. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
60. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347 (2003).
61. Bai, X. C., Fernandez, I. S., McMullan, G. & Scheres, S. H. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife* **2**, e00461 (2013).
62. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
63. Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
64. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
65. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
66. Kimanius, D., Forsberg, B. O., Scheres, S. H. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *Elife* **5**, e18722 (2016).
67. Chua, E. Y., Vasudevan, D., Davey, G. E., Wu, B. & Davey, C. A. The mechanics behind DNA sequence-dependent properties of the nucleosome. *Nucleic Acids Res.* **40**, 6338–6352 (2012).
68. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
69. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
70. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
71. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018).
72. Trabuco, L. G., Villa, E., Schreiner, E., Harrison, C. B. & Schulten, K. Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* **49**, 174–180 (2009).
73. Kidmose, R. T. et al. Namdinator - Automatic Molecular Dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *bioRxiv* 501197.
74. Lu, X. J. & Olson, W. K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* **31**, 5108–5121 (2003).
75. Lamichhane, R., Solem, A., Black, W. & Rueda, D. Single-molecule FRET of protein-nucleic acid and protein-protein complexes: surface passivation and immobilization. *Methods* **52**, 192–200 (2010).
76. Zhao, R. & Rueda, D. RNA folding dynamics by single-molecule fluorescence resonance energy transfer. *Methods* **49**, 112–117 (2009).

Acknowledgements

We thank Rishi Matadeen (formerly at NeCEN) for data collection of the NCP-intasome structure. We thank the EM and structural biology STP at the Crick for advice, computational and technical support. Histone plasmids were a gift from Joe Landry (Addgene) and 601 sequence was a gift from John Widom (Addgene). We are grateful to Pavel Afonine for help with Phenix real space refinement. This work was funded jointly by the Wellcome Trust, MRC and CRUK at the Francis Crick Institute (FC0010061, FC0010065). A.C. receives funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 820102). M.D.W. was funded by a Human frontiers Science Program long term Fellowship. The Single Molecule Imaging Group is funded by a core grant of the MRC-London Institute of Medical Sciences (UKRI MC-A658-5TY10), a Wellcome Trust Collaborative Grant (206292/Z/17/Z) and a Leverhulme Grant (RPG-2016-214).

Author contributions

A.C. and P.C. initiated the study. D.P.M. assembled the Intasome–NCP complex and LR determined the structure. D.P.M. performed pull-down assays. M.D.W. performed biochemistry, assembled NCP-D02-streptavidin and NCP-601 complexes and determined the structures. V.E.P. built all models into the EM maps. M.G. performed single molecule FRET experiments and data analysis, supervised by DSR. M.D.W., L.R. prepared and screened the cryo-EM grids and M.D.W. and A.N. collected cryo-EM data. M.D.W., P.C., and A.C. wrote the manuscript with input from the authors.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-019-12007-w>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Peer review information *Nature Communications* thanks Hitoshi Kurumizaka and other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



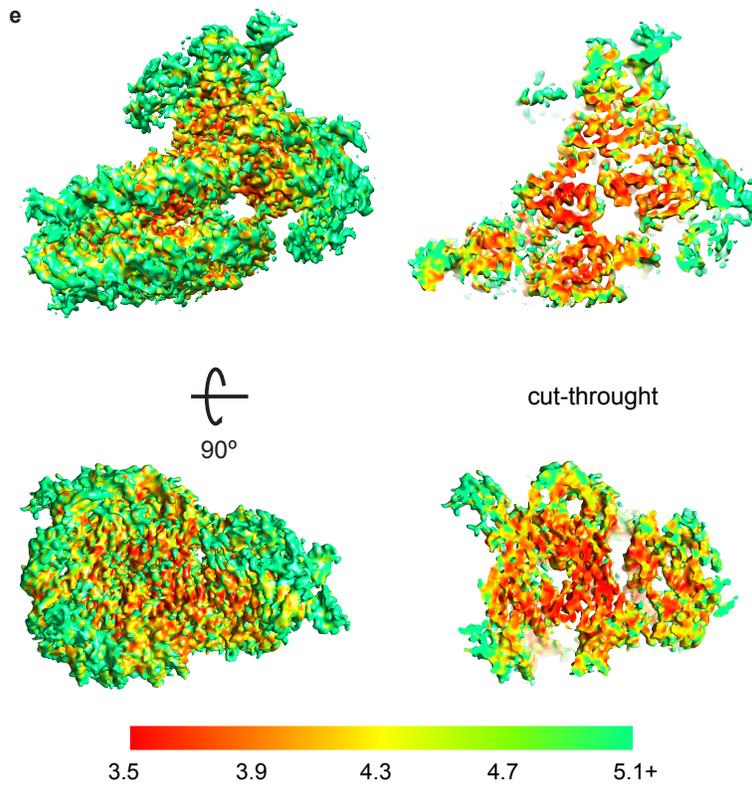
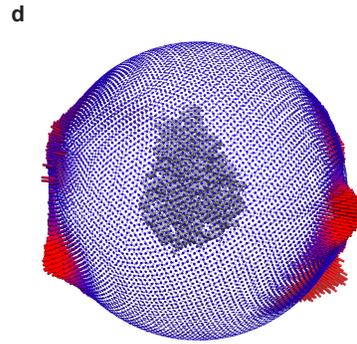
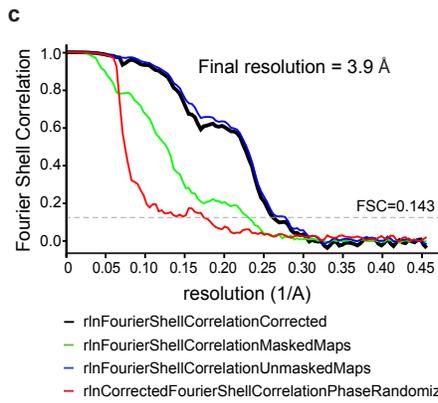
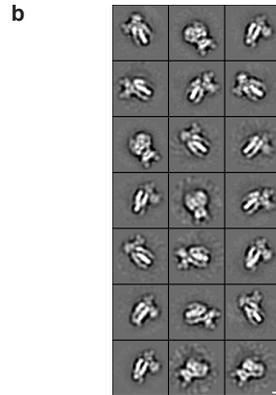
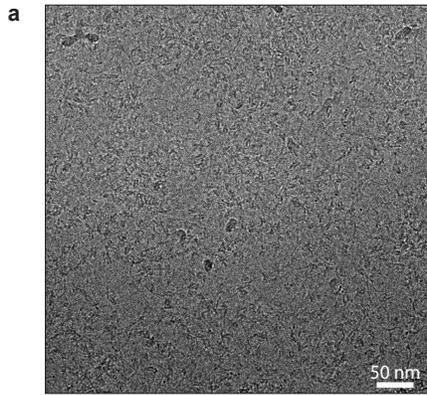
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

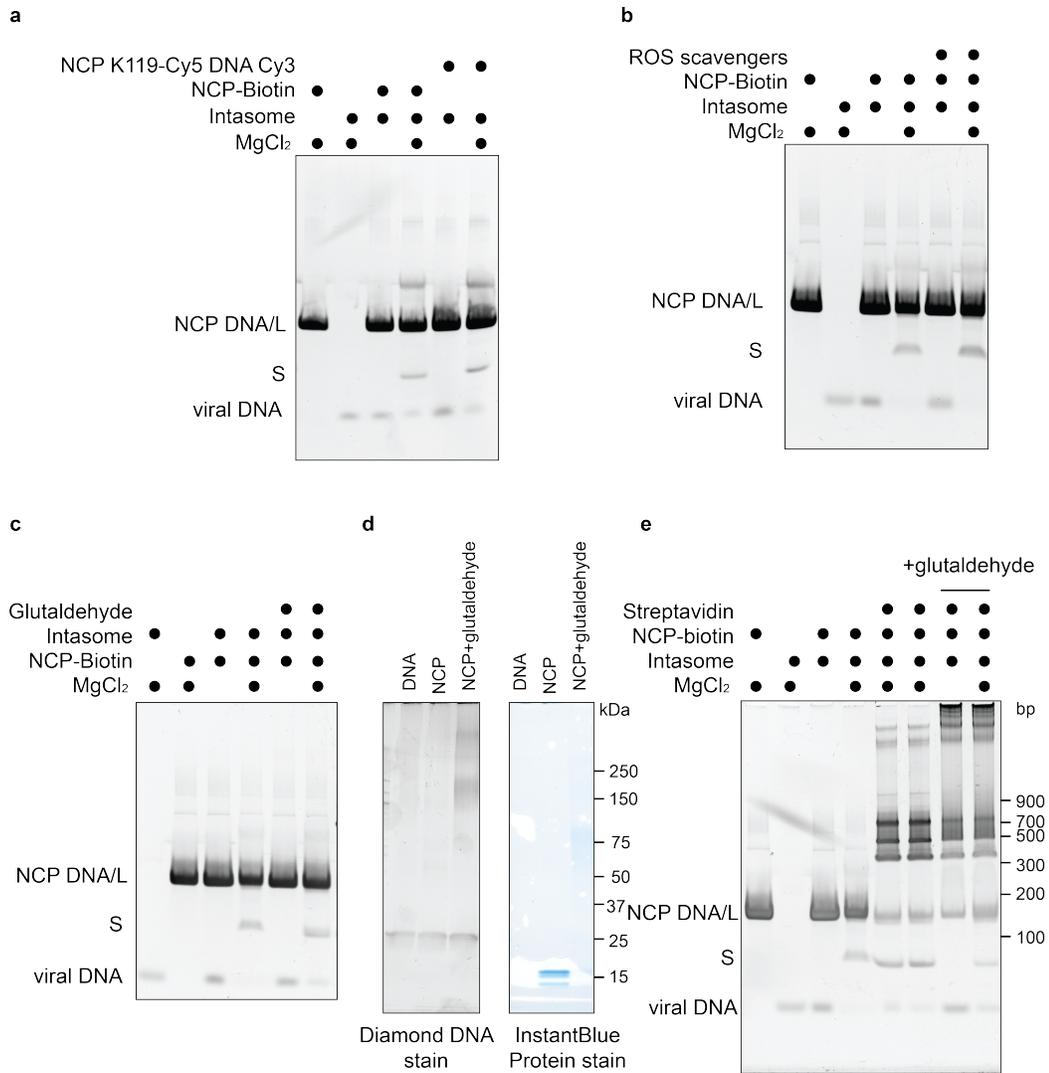
Supplementary Information

**Retroviral integration into nucleosomes through DNA looping
and sliding along the histone octamer**

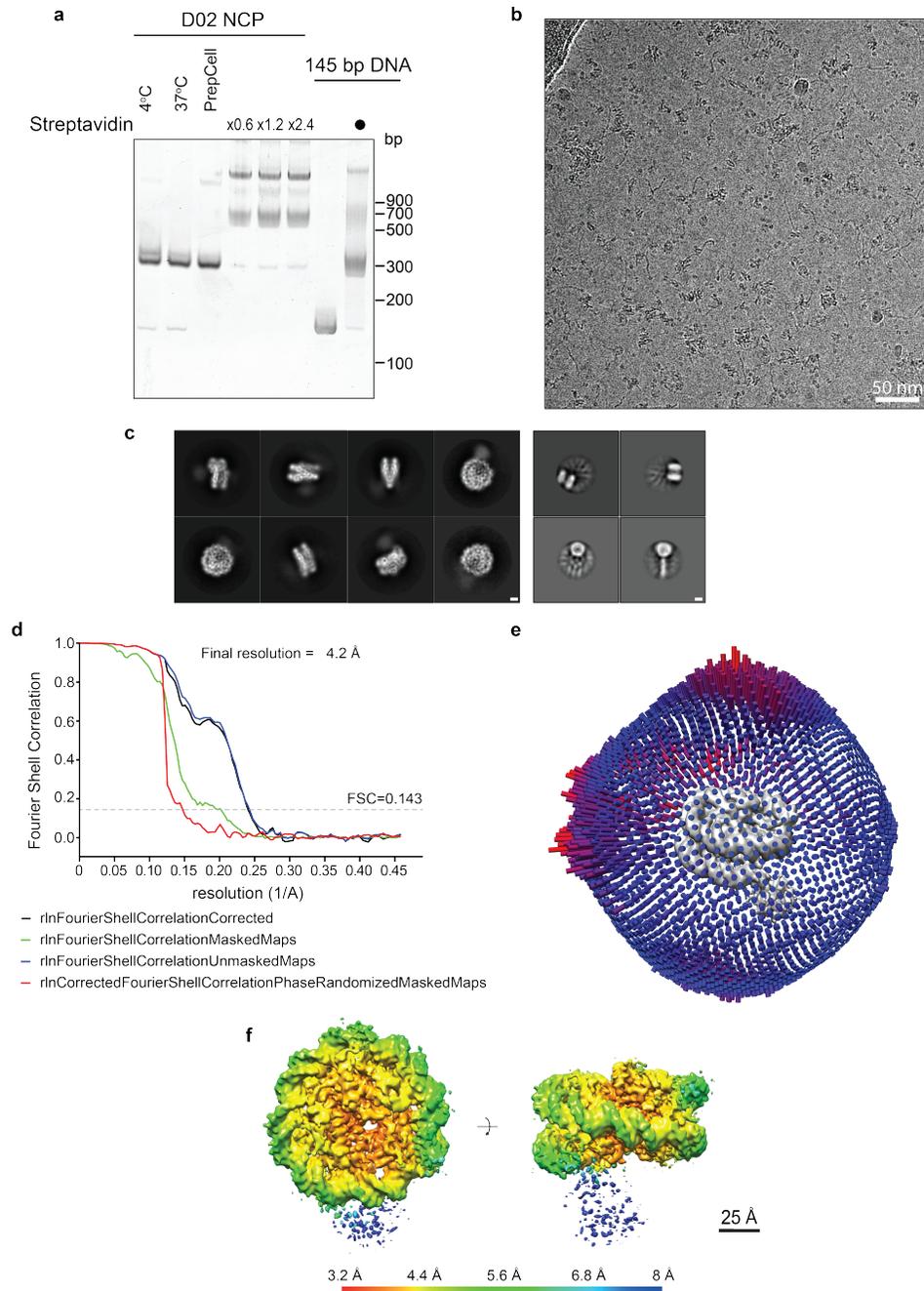
M. D. Wilson, L. Renault et al.



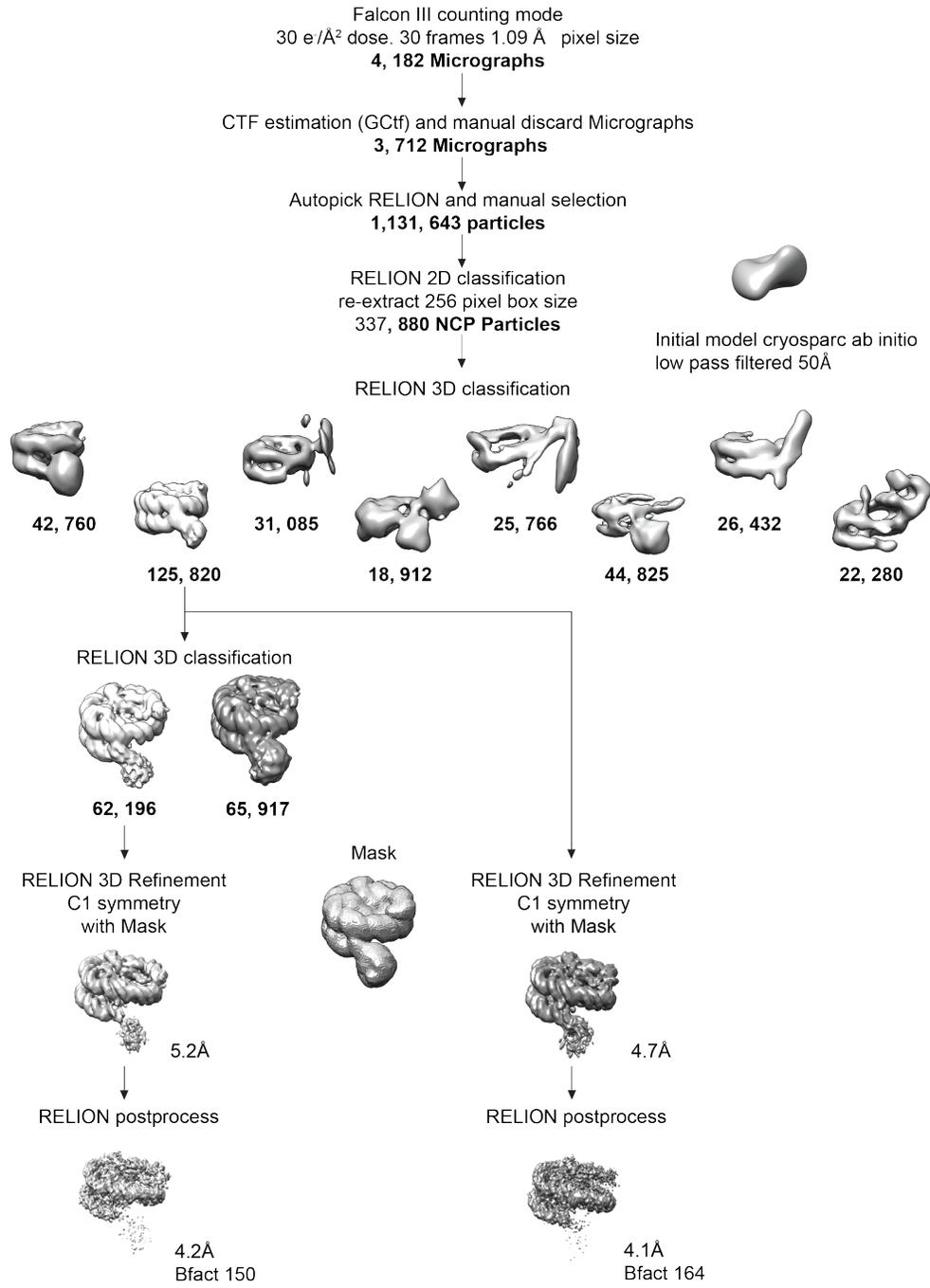
Supplementary Figure 1 Cryo-EM analysis of Intasome-D02 NCP. **(a)** Representative micrograph. **(b)** Representative 2D averages, white scale bar is 25 Angstrom long. **(c)** Gold standard FSC curve for the refined cryo-EM map. **(d)** Euler angle distribution plot for all particles included in the final map. Bar length and colour (blue low, red high) correspond to number of particles contributing to each view. **(e)** Cryo-EM map coloured according to local resolution estimated with ResMap.



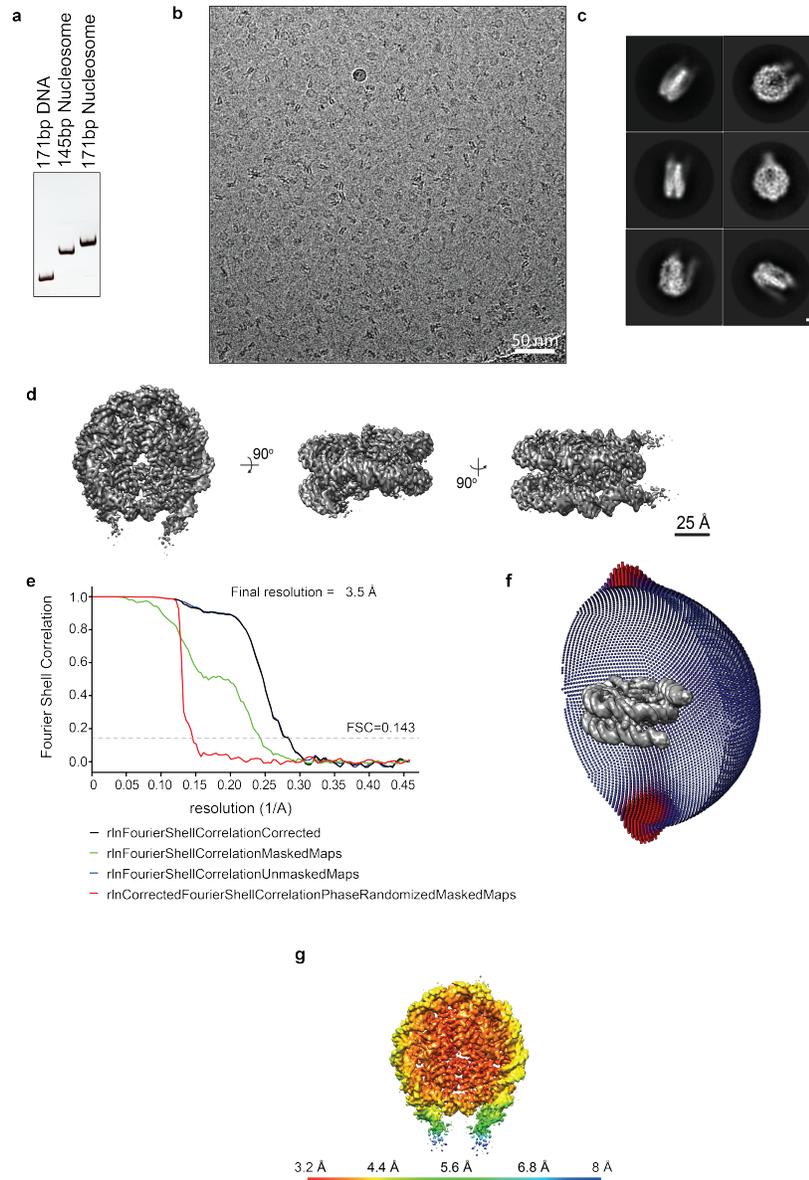
Supplementary Figure 2 Retroviral integration into various nucleosome derivatives. **(a)** Integration assay shows magnesium dependence of viral strand transfer. Strand transfer product S and L labeled. Assay performed with D02-NCP used for structure determination and fluorophore labeled nucleosomes used in single-molecule FRET experiments. **(b)** PFV Integration assay shows that ROS scavengers used in single-molecule FRET experiments do not affect the strand transfer reaction. **(c)** Glutaraldehyde cross-linking of D02 NCP does not prevent integration. **(d)** SDS-PAGE gels stained with Diamond nucleic acid stain (left) or InstantBlue protein stain (right). Glutaraldehyde cross-linked sample migrates as lower electrophoretic-mobility smear. **(e)** Efficient integration occurs with NCP-D02-biotin complexes pre-assembled with streptavidin, and cross-linked with glutaraldehyde.



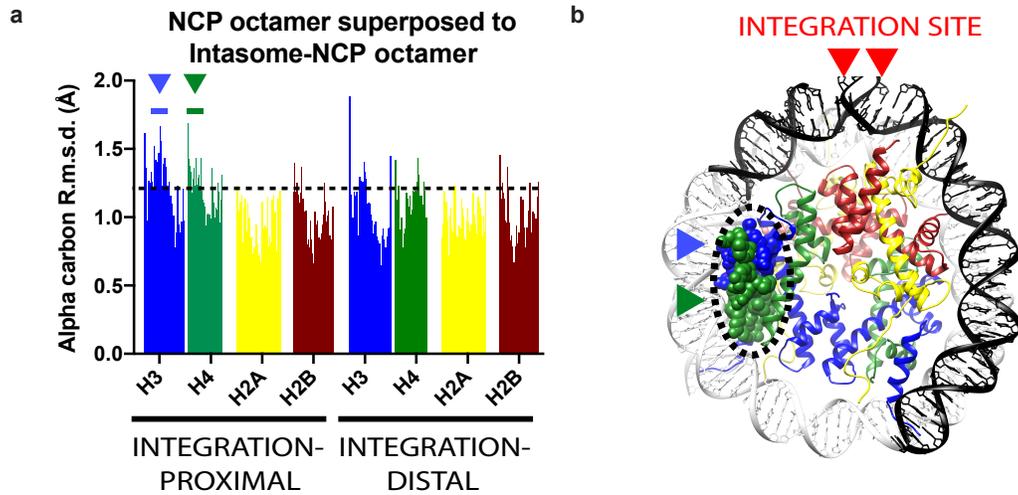
Supplementary Figure 3 D02 NCP-streptavidin preparation and cryo-EM characterisation. **(a)** Native gel electrophoresis of biotin-D02 wrapped NCP during purification and interaction with streptavidin. **(b)** Representative cryo-EM micrograph of NCP-D02-streptavidin. **(c)** Selected 2D averages of D02 NCP-streptavidin (left) and isolated or DNA-bound streptavidin (right), white scale bar is 25 Angstrom long. **(d)** Gold standard FSC curve for the final structure. **(e)** Euler angle distribution plot for all particles included in the final map. Bar length and colour (blue low, red high) correspond to number of particles contributing to each view. **(f)** NCP-D02-streptavidin coloured according to local resolution estimated in RELION. Streptavidin density is fragmented at the displayed threshold, due to the flexible tether with DNA.



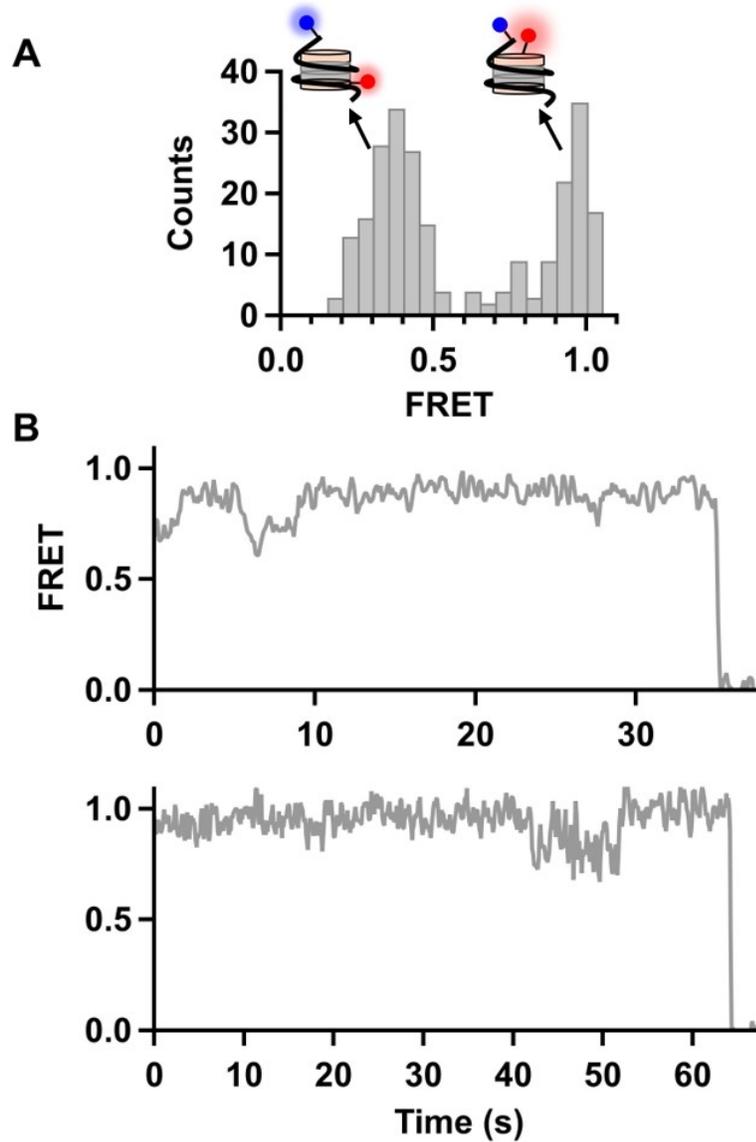
Supplementary Figure 4 Overview of cryo-EM single-particle reconstruction for D02 NCP-streptavidin.



Supplementary Figure 5 Electron microscopy of Widom 601 NCP. **(A)** Native gel electrophoresis of Widom 601 DNA isolated, and wrapped in a nucleosomes with either 145 base pair or 171 base pair-long DNA. **(B)** Representative cryo micrograph of nucleosome reconstituted with 171bp of Widom 601 DNA. **(C)** Selected representative 2D class averages, white scale bar is 25 Angstrom long. **(D)** Surface rendering of the NCP 171bp complex, shown in three views rotated about their axis. **(E)** Gold standard FSC curve for the refined cryo-EM map. **(F)** Euler angle distribution plot of all particles used for the final map (C2 symmetry imposed). Bar length and colour (blue low, red high) corresponds to number of particle images contributed to each view. **(G)** 171-base pair Widom 601 nucleosome, coloured according to local resolution estimated in RELION. Density for the DNA projecting away from the NCP core is visibly weaker.



Supplementary Figure 6 Intasome-induced octamer distortion. **(a)** After superposing Intasome-free and intasome-engaged D02 NCP via the histone octamers, the alpha carbon R.M.S.D. was measured for each histone chain. The most prominent intasome-induced movement is observed for histones H3 and H4 mapping on the nucleosome face proximal to the integration site. **(b)** Inspection of the integration-proximal face of the nucleosome shows that the most prominent movement clusters around DNA-interacting histone H3 and H4 elements, at a site downstream of histone H3 L1. Dotted line, blue and green triangles and atom spheres represent the region of most prominent alpha carbon movement. Red triangles indicate the integration site. Register-shifted DNA is black. Unperturbed DNA is white.



Supplementary Figure 7. Single-molecule FRET of D02 nucleosome core particles (A) Single-molecule FRET histogram of labelled nucleosomes (N = 245). Distal- and proximal-labelled nucleosomes yield 0.37 and 0.95 FRET-efficiency peaks, respectively. (B) Representative dynamic single-molecule FRET time trajectories of proximal-only fluorescent nucleosomes. Data were collected at 100 ms/frame and smoothed with a 3-point moving average.

mdw100	tggaacacaTGCACAGGATGTATATATCTGAC
mdw101	tggaacacaCCCTGGAGAATCCCGGTGC
mdw171	GGCTGTGTTTGTATCAAGTTACCTG
mdw166-2	/5Cy3/TGTCCAGGTTCTCCCTGTGGTAAAAACC
mdw155	TGTCCAGGTTCTCCCTGTGGTAAAAACC
mdw153	/5Biosg/GGCTGTGTTTGTATCAAGTTACCTG
mdw1	GTGATGGCGCTGCAGGAGGCTagcGAGGCCTACCTGGTGGGGCtatt
mdw2	GCCCCACCAGGTAGGCCTcgtAGCCTCCTGCAGCGCCATCACcgc
mdw3	GGTGGGGCTATTTGAGGATACCAACCTGgccGCCATTACGCCAAACGCGTCAC
mdw4	GTGACGCGTTTGGCGTGAATGGCggcCAGGTTGGTATCCTCAAATAGCCCCACC
mdw25	GAATATCCAGGCTGTTCTGCTGCCGAAAtgcACCGAATCTCACCACAAAGCGAAAGG
mdw26	CCTTTCGCTTTGTGGTGAGATTCCGGTgcaTTTCGGCAGCAGAACAGCCTGGATATTC
IO1	TGCGAAATTCATGACA
IO2	ATTGTCATGGAATTCGCA

Supplementary Table 1 List of oligos.

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

In this work by the Costa, Cherepanov, and Rueda groups, the authors present a 3.9 Å resolution cryo-EM structure of the prototype foamy virus intasome engaged with a nucleosome core particle. The structure, complemented with elegant single-molecule FRET experiments, immediately suggests a detailed mechanism for the structural changes on the nucleosome that are required for retroviral integration: nucleosomal DNA is shifted by approximately two base pairs past the histone core, with a concomitant small DNA loop lifted off the H2A/H2B surface.

This paper follows up on previous structural characterizations of the PVF intasome-nucleosome complex by Costa and Cherepanov. This earlier work had revealed an overall very similar architecture of the nucleosome-bound integration machinery. While it might therefore be argued that the structure presented in this current study is not fundamentally new per se, it is extremely important to stress that the previous structure was determined at substantially lower resolution, which prevented any detailed mechanistic understanding of retroviral integration into the nucleosome. In stark contrast, this current structure allows the authors to put forward a completely new loop-and-sliding mechanism for retroviral integration that is further corroborated by single-molecule FRET analysis from the Rueda lab.

The paper is well written and extremely clear and the experimental design is elegant. Particularly noteworthy is the clever use of a biotin moiety on the distal DNA arm to facilitate asymmetric particle alignment, resulting in an isolated D02 NCP structure of sufficient quality (highly challenging) to be compared to the intasome-NCP complex.

In sum, this well-executed and elegant study represents an important and timely advance that will be of substantial interest to a broad readership in various areas of chromatin and structural biology. I therefore enthusiastically recommend publication of this manuscript, essentially as is, at Nature Communications and only have very minor comments:

1) The authors utilize a FRET-based approach to detect conformational changes upon addition of intasome/Mg. The presence of two H2A copies per histone octamer gives rise to a heterogeneous labeling of the nucleosome with respect to Cy5. It would be good if the authors could include a FRET distribution for the labeled nucleosomes that shows both proximal and distal labeling configurations in the same histogram.

What was the labeling efficiency?

2) Addition of intasome/Mg gives rise to a second, lower-FRET peak, which is appropriately fit with a second Gaussian in Fig. 5F. The fit is not particularly convincing and could potentially be improved by increasing the sample size. I appreciate however that this may not be possible given the close proximity of the centroid positions.

3) Along the same lines, it is not clear why the distributions in C-E were fit with a sum of two Gaussians as in F.

4) The manuscript draws parallels between its loop-and-sliding mechanism and ATP-dependent chromatin remodeling. In light of this, the authors should consider extending their discussion to include recent single-molecule work (Sabantsev et al., Nature Communications 2019) that describes buffering of a small number of base pairs within the nucleosome as a salient feature of ATP-dependent chromatin remodeling, which may also be harnessed by other chromatin-interacting machinery.

Reviewer #2 (Remarks to the Author):

The authors have previously reported the cryo-EM structure of the precatalytic prototype foamy virus intasome in complex with a nucleosome core particle. The most notable feature of this structure was the looping out of DNA on the nucleosome surface at the site of intasome binding. This looping out is required to accommodate the sharp kink that target DNA adopts when engaged with intasome. However, the resolution of the previous structure was insufficient to reveal much detail beyond the lifting of DNA from the surface of the nucleosome at the site of intasome

binding. The new 3.9 Å structure of the postcatalytic complex reveals a much more detailed picture of the structure. As expected, the overall features are similar. The previous structure did not reveal how lifting of DNA from the nucleosome surface was accommodated, but the new structure, together with FRET data, shows that it is achieved by shifting the register of the flanking DNA on the nucleosome surface. It also reveals interactions of the nucleosome with the N-terminal histone H2A tail. This is high quality work that is well presented. However sliding of DNA on the nucleosome surface to accommodate the loop out is not surprising; the alternative of partial unwinding of flanking DNA would be energetically unfavorable. A more specialized journal may be more appropriate for publication of this work.

Minor comment:

Line 260: The suggested explanation for the requirement for magnesium to see a drop in FRET efficiency is unclear; the loop out (and by inference change in DNA register) is present prior to catalysis so why is the change in FRET only observed after catalysis?

Reviewer #3 (Remarks to the Author):

Retroviral integrase efficiently integrates the viral DNA into nucleosomes. A previous low resolution structural study by this collaborative team revealed that integrase directly contacts with the nucleosomal DNA by lifting DNA off the histone octamer surface at its integration site. To study conformational change of the nucleosomal DNA during vDNA integration process, in this study, the authors determined a cryo-EM structure of the prototype foamy virus intasome with a nucleosome at 3.9 Å resolution. The authors then found that the integrase induces two base-pair register shifting of the nucleosomal DNA, accompanying with the DNA lifting from the H2A-H2B surface. The DNA register shifting was also confirmed by the single-molecule Förster resonance energy transfer measurements. This is a high quality cryo-EM study, and interpretation of data is appropriate. Especially, it is an important progress that the authors clearly observed the cleavage site in the nucleosomal DNA. Therefore, I have only minor comments for improvement.

Comments:

The authors performed the structural analysis with Relion1.3 or 2.1. Relion3 may improve resolution better, although the conclusion of this work may not be affected. This is not essential issue.

In Figure S1d, cryo-EM map is missing in Euler angle distribution plot. please add it.

Please indicate scale of 2D class average in Figure 3B, Figure S1b, Figure S3c, and Figure S5c.

Scale bars in Figure S3f and Figure S5d represented as 25 nm. I suppose it should be angstrom. Please correct it.

What is the meaning of a caption "?" in Supplementary Fig2e, lane 5.

In the text, Line 246, I don't understand why the author cites Supplementary Figure 7. Please explain it or change the citation, if it is typo.

In Supplementary Fig6, it is helpful, if the authors show rmsd plots for Ca atoms for comparison between histone octamers, and discuss the relation between the octamer deformation and register shifting of the nucleosomal DNA.

Hitoshi Kurumizaka

We would like to thank the reviewers for comments that helped us improve our manuscript. Below we address the issues raised.

Reviewer 1:

We are thankful to this reviewer for very positive comments on our work. We are happy that this reviewer believes that our study “represents an important and timely advance”. Here is a point-by-point reply to the their comments:

1) The authors utilize a FRET-based approach to detect conformational changes upon addition of intasome/Mg. The presence of two H2A copies per histone octamer gives rise to a heterogenous labeling of the nucleosome with respect to Cy5. It would be good if the authors could include a FRET distribution for the labeled nucleosomes that shows both proximal and distal labeling configurations in the same histogram.

As noted by the reviewer, we observe a mixed population of nucleosomes with tails labelled on either copy of H2A in the octamer. As a consequence, we detect markedly distinct FRET efficiency signals, depending on whether H2A tails proximal or distal from the labelled DNA are measured. As requested by the reviewer, in Supplementary Figure 7 we have added a FRET distribution for the labelled nucleosomes, showing both proximal and distal labelling configurations in the same histogram.

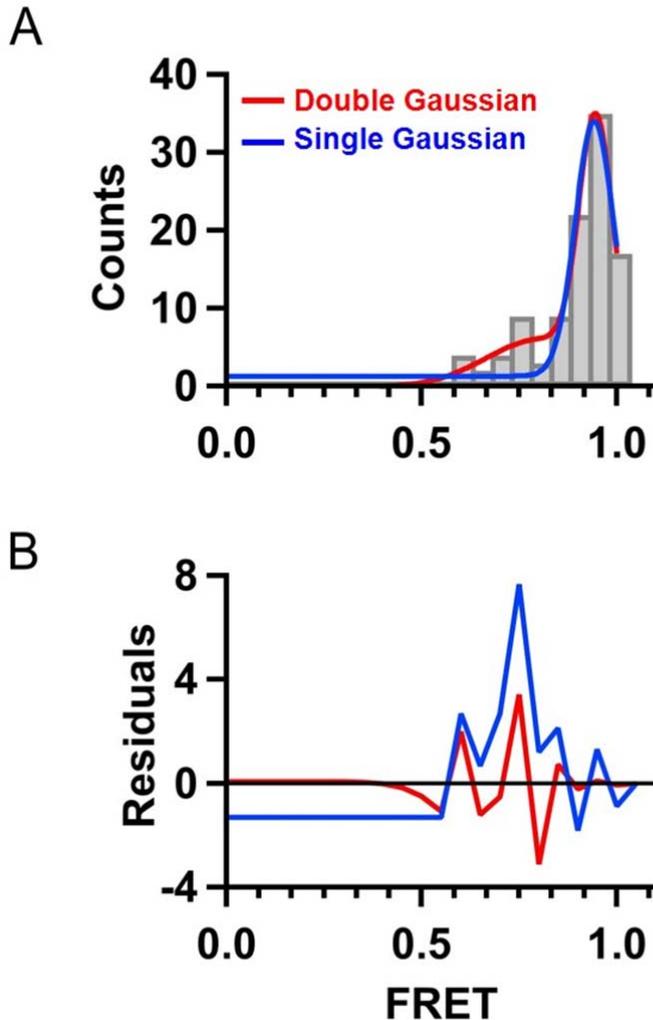
What was the labeling efficiency?

K119C labelling efficiency was 68%, as established by spectrophotometric (A280 vs A650) measurements, and correct labelling was confirmed by intact mass spectrometry. We have now modified the main methods section to include this information. Assuming a normal distribution, we expect the majority of nucleosomes to contain a single label on H2A K119C. Because of the experimental design, non-labelled nucleosomes do not produce a FRET signal. To ensure that nucleosomes labelled on both K119C residues were excluded from our measurements, we only quantified FRET traces where Cy3 was seen to bleach to background levels, during the course of the experiment.

2) Addition of intasome/Mg gives rise to a second, lower-FRET peak, which is appropriately fit with a second Gaussian in Fig. 5F. The fit is not particularly convincing and could potentially be improved by increasing the sample size. I appreciate however that this may not be possible given the close proximity of the centroid positions.

3) Along the same lines, it is not clear why the distributions in C-E were fit with a sum of two Gaussians as in F.

We have done one of the fits with a double and a single Gaussian distribution:



The figure clearly shows that the single distribution fit cannot represent the small shoulder at lower FRET values (smaller, randomly distributed residuals - blue curve). Based on this evidence, we think the double distribution is warranted. The observed shoulder derives from a small fraction of dynamic molecules that transiently visit lower FRET states (see **Sup. Fig. 7**), thereby lowering the average FRET value. This shoulder is different, however, from the large population observed at lower FRET that arises in the presence of IN/Mg²⁺ (**Figure 5F**). With the latter conditions in fact, we reproducibly detect a significant population of molecules featuring a ~0.15 decrease in FRET efficiency. We agree with the reviewer that additional data will not improve the distribution, given the close proximity of the centroid positions. Based on our experience with a very similar experimental design (Wilhoft and Ghoneim *Science*, 2019), we are confident that the change in FRET observed with the addition of IN/Mg²⁺ reports on an alteration in nucleosomal register. For clarity, we would prefer not to change the fit in the figures.

4) The manuscript draws parallels between its loop-and-sliding mechanism and ATP-

dependent chromatin remodeling. In light of this, the authors should consider extending their discussion to include recent single-molecule work (Sabantsev et al., Nature Communications 2019) that describes buffering of a small number of base pairs within the nucleosome as a salient feature of ATP-dependent chromatin remodeling, which may also be harnessed by other chromatin-interacting machinery.

We thank the reviewer for this important remark. We now reference the Sabantsev paper in the main text.

Reviewer 2:

We thank this reviewer for noting that our new 3.9 Å resolution reconstruction of the postcatalytic nucleosome-intasome “reveals a much more detailed picture” of the complex. The reviewer acknowledges that the “new structure, together with FRET data” reveals that nucleosomal DNA looping essential for integration “is achieved by shifting the register of the flanking DNA on the nucleosome surface”. Finally, the reviewer acknowledges our biochemical and structural analysis revealing “interactions of the nucleosome with the N-terminal histone H2A tail”. We would like to thank this reviewer for stating that “this is high quality work that is well presented”.

However sliding of DNA on the nucleosome surface to accommodate the loop out is not surprising; the alternative of partial unwinding of flanking DNA would be energetically unfavorable.

We now state in the introduction that underwinding of flanking DNA could be one alternative to nucleosomal DNA repositioning. DNA binding proteins such as SSB can melt DNA in the absence of any ATPase activity. It is therefore unclear to us why this reviewer states that nucleosomal DNA repositioning should be energetically more favourable than DNA underwinding. Notably, only one case has been reported so far (which was published after our first original submission), where nucleosomal DNA repositioning occurs upon binding of an interactor that lacks ATPase activity. We understand this observation was deemed noteworthy, given that it featured in the title of a recent article on UV-DDB bound to nucleosomes bearing damaged DNA (Matsumoto et al Nature 2019). This study is now cited in our main text. We are likewise excited about our discovery that IN induces nucleosomal DNA repositioning to allow DNA access to the catalytic site for integration.

A more specialized journal may be more approximate for publication of this work.

It should be noted that Reviewer 1 states that our manuscript describes “completely new loop-and-sliding mechanism for retroviral integration”. In Reviewer 3’s opinion, observing a cleavage site in the nucleosomal DNA represents “important progress”. Thus, we respectfully disagree with Reviewer 2 and strongly agree with Reviewers 1 and 3 about the importance of our work.

Minor comment:

Line 260: The suggested explanation for the requirement for magnesium to see a

drop in FRET efficiency is unclear; the loop out (and by inference change in DNA register) is present prior to catalysis so why is the change in FRET only observed after catalysis?

We have expanded on our discussion regarding magnesium requirement to observe a drop in FRET efficiency. As pointed out by the reviewer, according to crystallographic work, the degree of bending required for capturing target DNA matches that required for strand transfer. Likewise, in the pre-assembled intasome-NCP complex, the nucleosomal DNA path pre- and post-integration appears virtually identical. In our single-molecule experiment, however, a change in FRET was only observed in the presence of magnesium that promotes integration. This is expected given that our FRET assay was performed by flowing intasome on a glass slide with tethered nucleosomes, and not using a pre-assembled intasome-NCP complex. Single-molecule work by Fishel and Yoder, performed under comparable experimental conditions (though using non-chromatinised DNA), showed that Intasome spends most of the time scanning DNA by one-dimensional diffusion, before suitable DNA segments are recognised for strand transfer (Jones et al. *Nat comm*, 2016). When magnesium is omitted in our experiment, we therefore expect that intasome will spend most of the time scanning nucleosomal DNA. This search is unlikely to result in any DNA deformation. As the looped & register-shifted state is expected to accumulate upon integration, it is therefore not surprising that the nucleosomal DNA sliding can be clearly detected in the presence of magnesium, which is essential for catalysis.

As suggested by reviewer 1, an alternative explanation is that DNA perturbation and sliding could be functionally distinct steps. This could be mediated by histone buffering of DNA displacement, as observed previously³³. In this model tension in the lifted DNA at the integration site could be accommodated by protein-DNA contacts within the target capture complex, without a concomitant shift in DNA, which only occurs upon full strand transfer catalysed upon addition of magnesium.

Reviewer 3

We thank this reviewer for stating that our study is “high quality” and “appropriate” data interpretation, and that observing the cleavage site in the nucleosomal DNA is deemed “important progress”.

Comments:

The authors performed the structural analysis with Relion1.3 or 2.1. Relion3 may improve resolution better, although the conclusion of this work may not be affected. This is not essential issue.

We agree and attempted reprocessing one of our datasets (Widom 601 NCP) with Relion3 and obtained no improvement, including using new implementations of CTF refinement and Bayesian polishing. Given that all conclusions of our study are solidly supported by the structures presented at the current resolution, and validated by

single-molecule FRET and mutagenesis studies, we believe that reprocessing our data with more recent releases of the same software package is not necessary in this instance.

In Figure S1d, cryo-EM map is missing in Euler angle distribution plot. please add it.

We have now added the Euler plot in Figure S1d.

Please indicate scale of 2D class average in Figure 3B, Figure S1b, Figure S3c, and Figure S5c.

We have now added scale bar for 2D averages in the captions of Figures 3B, S1B, S3C and S5C.

Scale bars in Figure S3f and Figure S5d represented as 25 nm. I suppose it should be angstrom. Please correct it.

We would like to thank the referee for spotting this mistake. This has now been corrected.

What is the meaning of a caption “?” in Supplementary Fig2e, lane 5.

This was a mistake, thanks for spotting it. Lane 5 in Supplementary Figure 2e should be labelled Mg^{2+} . Now corrected.

In the text, Line 246, I don't understand why the author cites Supplementary Figure 7. Please explain it or change the citation, if it is typo.

Thanks for spotting this. Supplementary Figure 7 should not be cited here and has been removed.

In Supplementary Fig6, it is helpful, if the authors show rmsd plots for Ca atoms for comparison between histone octamers, and discuss the relation between the octamer deformation and register shifting of the nucleosomal DNA.

This is a great suggestion. In Supplementary Figure 6, we have now added the RMSD plots for Ca atoms comparing histone octamers in the strand transfer complex and isolated D02 nucleosome. In the Discussion, we now examine the relation between octamer deformation and nucleosomal DNA shifting.