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

Marcus D. Wilson[1,§,%], Ludovic Renault[1,§,#], Daniel P. Maskell[2,&], Mohamed Ghoneim[3], Valerie E. Pye[2], Andrea Nans[4], David S. Rueda[3,5]*, Peter Cherepanov[2,]*, Alessandro Costa[1,]*

[1]Macromolecular Machines Laboratory, The Francis Crick Institute, NW1 1AT, London, UK

[2]Chromatin structure and mobile DNA Laboratory, The Francis Crick Institute, NW1 1AT, London, UK

[3]Single Molecule Imaging Laboratory, MRC London Institute for Medical Science, W12 0NN, London, UK

[4]Structural Biology Science Technology Platform, The Francis Crick Institute, NW1 1AT, London, UK.

[5]Molecular Virology, Department of Medicine, Imperial College London, W12 0NN, London, UK

[%]Present address: Wellcome Centre for Cell Biology, University of Edinburgh, EH9 3JR Edinburgh, UK

[#]Present address: NeCEN, University of Leiden, 2333CC, Leiden, The Netherlands

[&]Present address: Faculty of Biological Sciences, LS2 9JT, Leeds, UK

[§]Equal contribution.

*Correspondence: alessandro.costa@crick.ac.uk, peter.cherepanov@crick.ac.uk and david.rueda@imperial.ac.uk.

Abstract

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 with chromatin remodelers.

Introduction

Integration of the reverse-transcribed retroviral genome into a host-cell chromosome is catalysed by integrase (IN), an essential viral enzyme (reviewed in [1]). 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 DNA targets, DNA bending 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 intasomal 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 resolution[15]. 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 DNA [15]. 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 FRET (smFRET) 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 new findings we uncover unexpected similarities between the mechanisms of retroviral integration and ATP-dependent chromatin remodeling [21-23].

Results

Structure of Intasome-NCP strand-transfer complex

To understand intasome strand transfer into NCPs, we assembled the complex of the PFV intasome and the NCP containing a native human DNA sequence (termed D02), selected for its ability to form a stable PFV-NCP complex[15]. Following isolation by size exclusion chromatography, the intasome-NCP complex was incubated in the presence of Mg^{2+} to facilitate strand transfer[15]. 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 Figure 1** and **Supplementary Table 1**). Docking known crystallographic coordinates into the cryo-EM map and use of 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] (**Figure 1**). The new structure is overall similar to the original lower-resolution intasome-NCP complex, which was captured in the pre-catalytic state (**Figure 1A**), confirming that strand transfer is not accompanied by large conformational rearrangements[6]. 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 structure of the PFV strand transfer complex and the lower-resolution intasome-NCP cryo-EM data.

Local resolution is highest (~3.5 Å) for the histone octamer core, while DNA density on the outer perimeter tends to decrease (~4-4.5 Å), as observed previously for other NCP structures (**Supplementary Figure 1**) [25-27]. 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 higher local resolution (~3.7 Å). Notably, a discontinuity in the cryo-EM density resulting from the

nucleosomal DNA cleavage at the site of integration (**Figure 1B**) confirms that strand transfer has indeed occurred in our nucleoprotein assembly as observed biochemically[15] (**Supplementary Figure 2**).

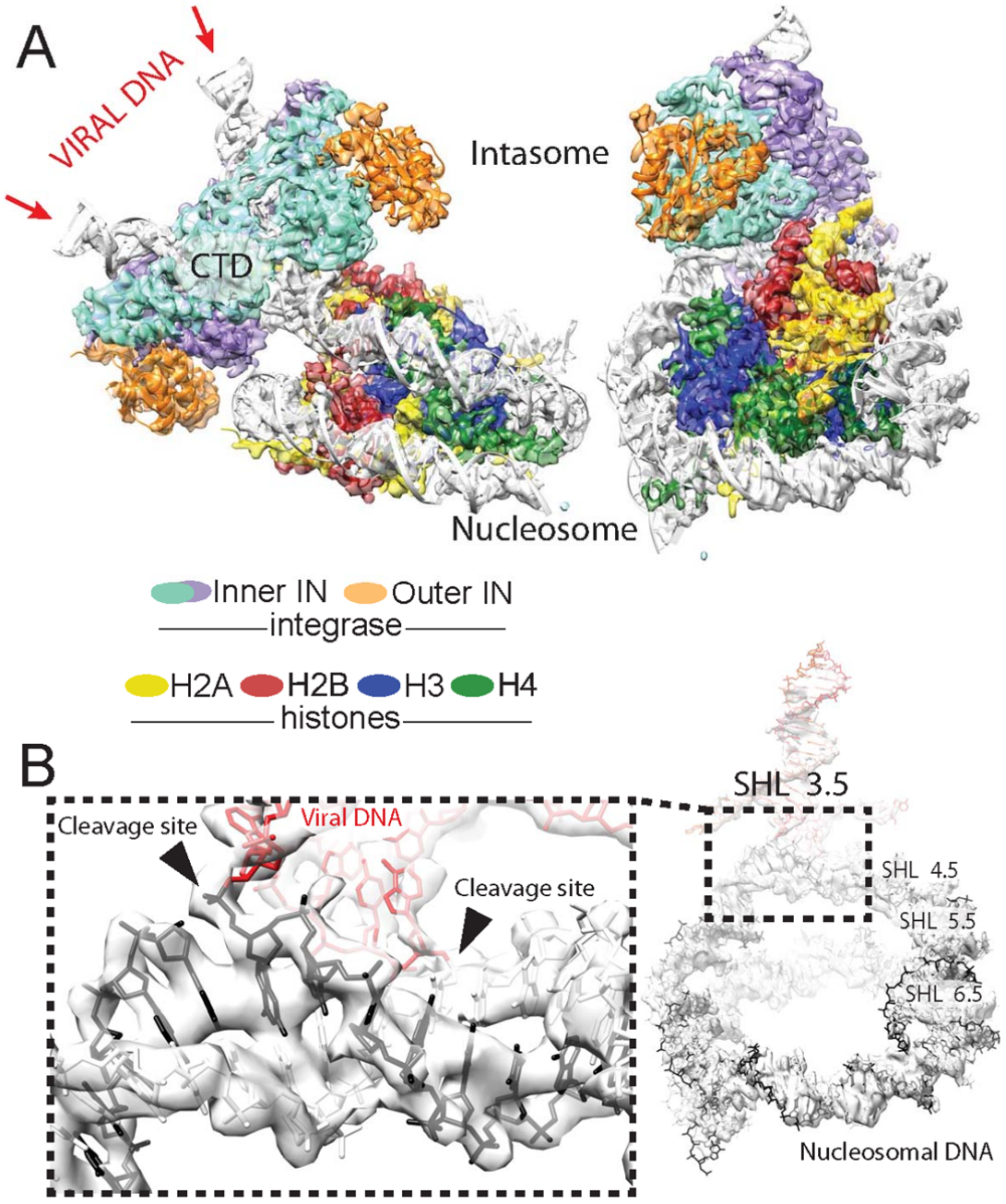

Figure 1 Intasome-NCP strand-transfer complex visualised by cryo-EM. **(A)** 3.9 Å resolution structure of the post-catalytic intasome-NCP complex. **(B)** Covalently linked viral (red) and nucleosomal (black) DNA. Integration occurs at SHL location 3.5.

Table 1 Data collection and processing information

119

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 ab initio	CryoSPARC ab initio
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	240x240x240	256 x 256 x 256	256 x 256 x 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	
Ramachandaran favoured/outlier	94.3% / 0%	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	
Molprobity score	1.91	1.45	
PDB ID	PDB: 6RNY	PDB: 6R0C	

120
121

122 Intasome engages nucleosomal DNA non-symmetrically at two distinct sites: at the strand
123 transfer site as well as at the opposing gyre, which nestles in the cleft between one catalytic

and one outer IN subunit (**Figure 1A**). Near the integration site, the alpha C-helix of histone H2B makes direct contact with the C-terminal domain of one catalytically competent IN subunit, providing corroborating evidence for the reported role of IN residues Pro135, Pro239 and Thr240 in engaging C-terminal H2B [15]. 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 close proximity between positively charged Lys-9 and Arg-11 on H2A and the catalytically competent IN C-terminal domain (**Figure 2A**). Coherently, we observe that complex formation is reduced with an NCP containing a histone H2A truncation of the first 12, but not the first 8 residues (**Figure 2B**). Furthermore, Ala substitutions of either H2A at Lys-9 or Arg-11 affect complex stability, while a combination of the two substitutions fully abrogates the stable complex formation under conditions of the pull-down assay **(Figure 2C)**.

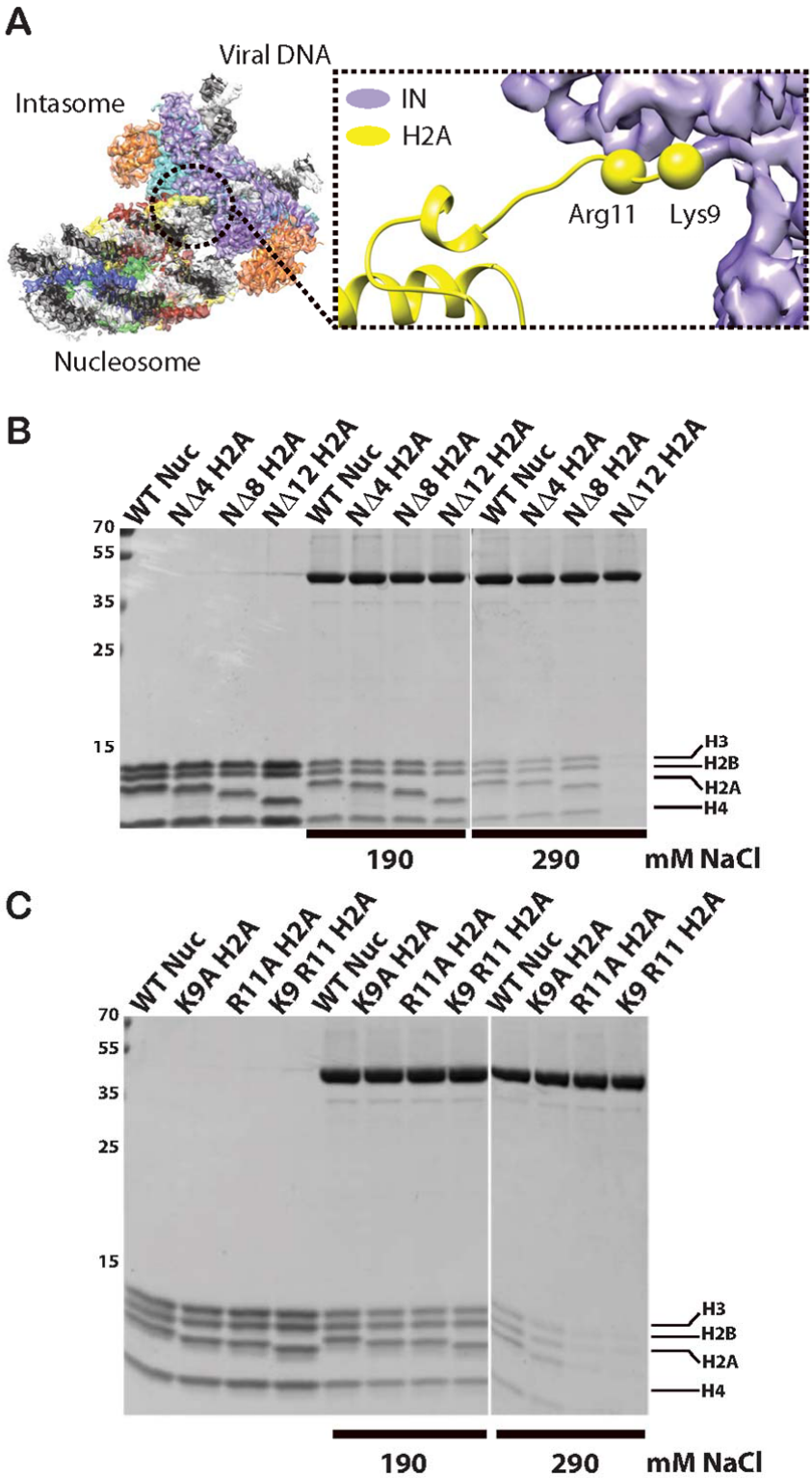

Figure 2 Intasome interaction with the N-terminal histone H2A tail. (**A)** Histone H2A residues Lys-9 and Arg-11 play a key role in intasome-NCP interaction. **(B)** Pull-down assay with immobilised intasome binding to NCP H2A tail deletion variants. (**C)** Pull-down assays with immobilised intasome binding to NCP H2A N-terminal tail variants containing single- or double-point mutations.

Asymmetric reconstruction of isolated NCP bearing the human D02 DNA sequence

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 chromatin remodelers such as SWR1. Interestingly, DNA looping by SWR1 is accompanied by both sliding of nucleosomal DNA, as well as histone octamer distortion [22]. 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 and contrast 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 sequence[15].

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[15,28]. 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 (data not shown). 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 Figure 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 NCP[15]. 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 (**Figure 3A**). Critically, streptavidin attachment did not affect NCP stability, nor the ability of intasome to integrate into NCPs (**Supplementary Figure 2**). Crosslinked D02 NCPs, imaged by cryo-EM and analysed by two-dimensional (2D) averaging, revealed multiple views of the coin-shaped NCP assemblies (**Figure 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 Figure 3**). Next, we used single-particle reconstruction to determine the 4.2 Å resolution structure of NCP-D02-streptavidin complex (**Supplementary Figure 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 any stabilising interaction with the

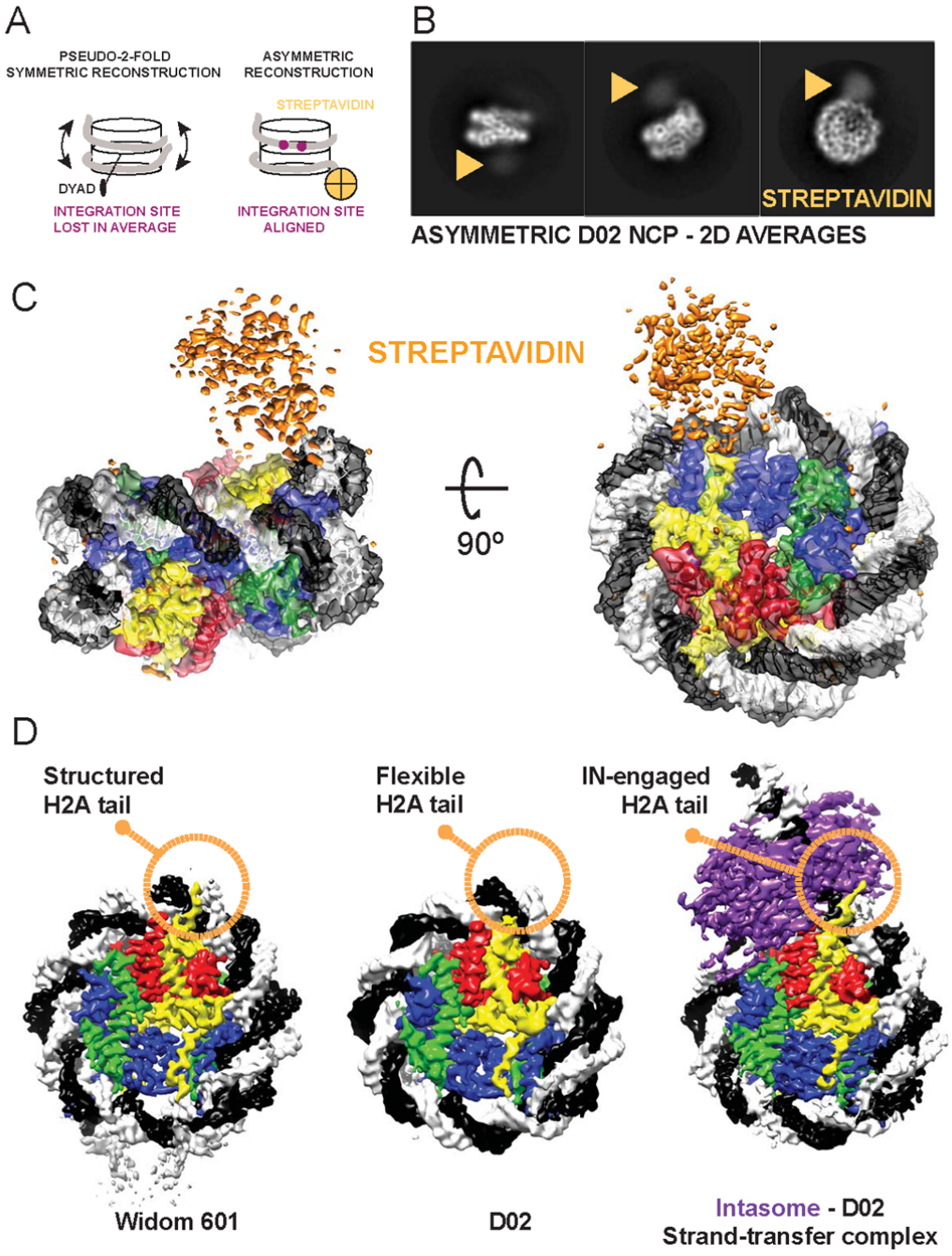

Figure 3 Asymmetric reconstruction of the isolated D02 NCP. **(A)** Streptavidin labelling allows asymmetric reconstruction of the D02 NCP, which avoids pseudo-two-fold symmetry averaging of the integration-site DNA. **(B)** 2D averages of labelled D02 NCP particles reveal a discernible, diffuse density 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 Figure 2B and C).

NCP core (**Figure 3C** and **Supplementary Figure 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 complex[15]. 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 Figure 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 is a poor substrate for intasome binding and integration. To this end, we solved the 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 Figure 5**). Only linker DNA fragments display a degree of flexibility in the Widom 601 structure. We postulated at this stage that flexible NCP arms in D02 might favour nucleosomal DNA repositioning required to support DNA looping, prompting us to further investigate the mechanism.

A second notable feature in the D02 NCP structure 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 [7,29-31]. We speculate that loose DNA-engagement renders the histone H2A tail available for intasome binding as observed in our strand-transfer complex, hence improving substrate selection (**Figure 3D**).

Retroviral integration causes a shift in 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 with a global r.m.s. deviation of backbone atom positions of 2.7 Å (**Supplementary Figure 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 (**Figure 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 (**Figure 4B** and **Supplementary Movie 1**).

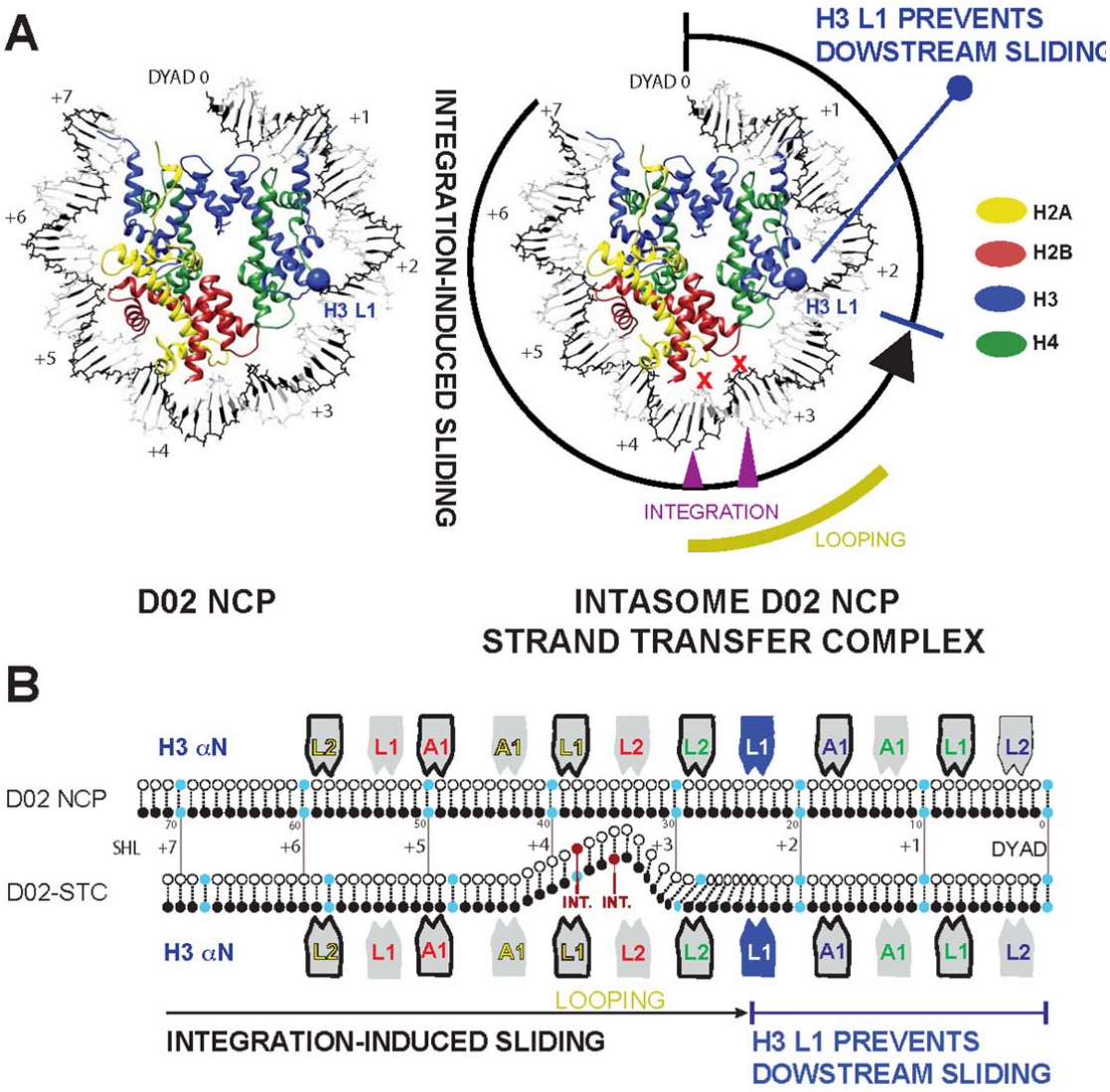

Figure 4 Integration-promoted DNA sliding observed by cryo-EM. **(A)** On the left, slabbed view of the isolated D02 NCP. Histone H3 L1 element is highlighted with a blue ball. On the right, DNA looping required for retroviral integration causes a shift in the DNA register, which extends from SHL 7 to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

To validate the DNA-register change observed in our structural models, we turned to a single molecule Förster resonance energy transfer (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 (**Figure 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

236 Mg^{2+} (**Figure 5B**). In reconstituted NCPs, single H2A labels were found either proximal to, or
237 distal from, the Cy3-modified DNA end. The main energy transfer group deriving from the
238 proximal fluorophore pair centred around 0.95 FRET efficiency, while the second distal
239 fluorophore pair peak centred around 0.37 transfer efficiency (**Supplementary Figure 7A**).
240 We focused our analysis on the 0.95 FRET group, as any shift in nucleosomal DNA register
241 would cause more pronounced changes in FRET efficiency in this population. In all tested
242 conditions, FRET efficiency was stable, with a minor population (~10%) of traces exhibiting
243 slight changes in FRET intensity (**Figure 5C and Supplementary Figure 7B**).
244 Supplementing the NCP with intasome or Mg^{2+} did not result in any significant FRET change
245 (**Figure 5D** and **5E**). However, when strand transfer was induced by adding both intasome
246 and Mg^{2+} (**Supplementary Figure 7**), a separate, ~0.8 FRET population appeared (**Figure**
247 **5F**). This second population is consistent with a shift in register of the DNA moving away
248 from the K119C-Cy5 H2A residue (**Figure 5A**). These data are in good agreement with our
249 comparative cryo-EM analysis indicating that intasome-mediated looping required for
250 integration promotes sliding of nucleosomal DNA (**Figure 4A**). In fact, the observed drop in
251 FRET efficiency indicates a small but significant shift in the DNA register that corresponds to
252 less than 4 bp, according to a calibration previously obtained with Widom 601 NCPs[22].
253 Crystallographic and cryo-EM structures of pre-catalytic assemblies of intasome bound to
254 DNA or nucleosomes established that target capture alone leads to DNA bending and
255 nucleosomal DNA remodelling[15,16]. The new post-catalytic intasome-NCP structure reported
256 here confirms that no change of DNA looping occurs at the integration site after strand
257 transfer. However, in our single molecule experiments where intasome was added to
258 nucleosomes tethered to glass slides, a drop in FRET efficiency was only observed in the
259 presence of magnesium required for integration. Although this observation was surprising to
260 us, it is reasonable to infer that the increased stability obtained with covalent link formation
261 yields consistent FRET efficiency-drop derived from intasome binding and nucleosomal DNA
262 repositioning.

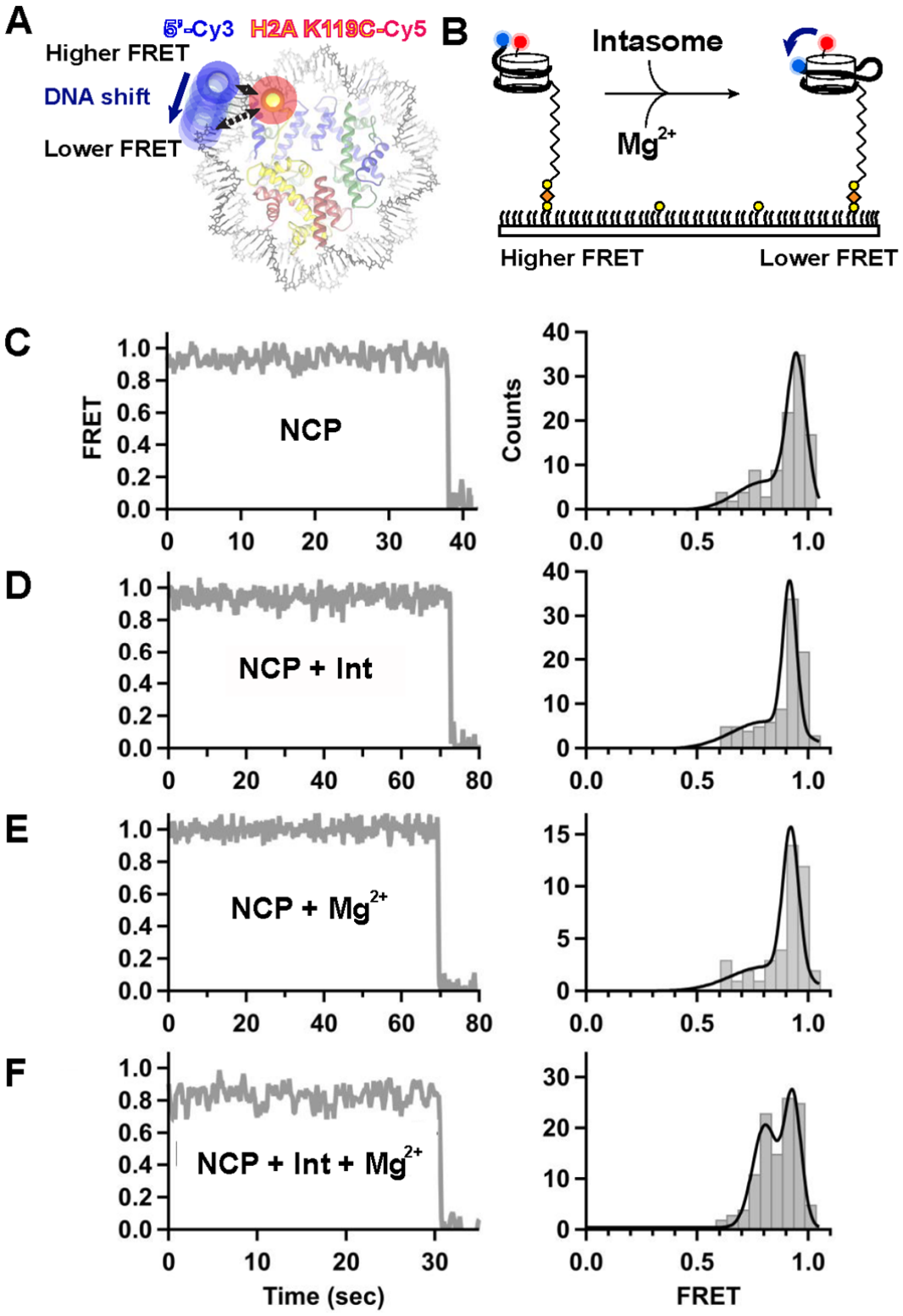

Figure 5 Integration-promoted nucleosomal DNA sliding observed by single molecule FRET. (**A**) Fluorescently labelled NCP: labelled octamer (**Cy5** on H2A, red) wrapped with biotinylated labelled DNA (**Cy3** on exit site, blue). **(B)** NCPs are surface-immobilised on neutravidin (orange) coated biotin-PEG (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 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.

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 helix[32], and how common docking sites on the histone octamer are recognised by different interactors[33-37]. Over the last four years, cryo-EM has started to provide a dynamic view of the NCP [25,38-41]. 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 lattice[42]. 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 dimer[27]. Spectacular views of progressively unwrapped NCPs have been obtained for transcribing RNA polymerase II captured during NCP passage [43,44]. Moreover, cryo-EM provided the first glimpses of ATP-dependent NCP translocation through a mechanism involving DNA looping and sliding along the histone octamer [22,38,45-51].

Our high-resolution view of a post-catalytic intasome-NCP complex provides a new example of a local remodeling of nucleosomal DNA. Although previous work established formation of a DNA loop during productive intasome-NCP interaction, it was not clear whether the loop forms by local stretching of the double helix 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 NCP remodelers. 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 site[15], 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 (**Figure 4B** and **Supplementary Movie 1**).

Using cryo-EM, Kurumizaka and colleagues have recently shown that the same H3 L1-DNA interaction stalls RNA polymerase II during NCP passage[43]. 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 [22,45,52]. It is tempting to speculate that the concerted action of IN and SHL 2 remodelers could act synergistically during DNA unpeeling and strand-transfer complex disassembly, required to complete retroviral integration. In support of this notion, a functional coupling during retroviral integration has been reported between HIV-1 integrase and Snf2-related SWI/SNF chromatin remodelling complex [53].

Supplementary Movie 1: Nucleosome core particle morphed between the isolated and the strand transfer complex state.

Materials and 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 20 mM imidazole. Bound proteins were eluted with 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 concentrated to 10 mg/ml, as estimated by spectrophotometry at 280 nm and stored at -80^{o}C.

To assemble the intasome a mixture containing 120 μM PFV IN and 20 μM pre-annealed DNA oligonucleotides 5'-TGCGAAATTCCATGACA and 5'-ATTGTCATGGAATTTCGCA (IDT) in 500 mM NaCl was dualized against 50mM BisTris propane-HCl pH 7.45, 200 mM NaCl, 40 μM $ZnCl_2$, 2 mM DTT for 16 h at 18°C. 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 [15,54]. 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 x 100 µl) were pooled, filtered and purified on a ResorceQ column as described [15]. NCPs were assembled by salt dialysis as described [15,29,54] and heat repositioned at 37°C for 30 minutes. D02 containing NCPs were further purified using a PrepCell apparatus with a 5% polyacrylamide gel (BioRad).

NCP-streptavidin complex formation

Purified *Streptomyces avindii* streptavidin powder (Sigma-Aldrich) was resuspended in 20 mM HEPES-NaOH pH 7.5, 150 mM NaCl at a final concentration of 35 µM. 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 µM) was incubated with 0.3 µM streptavidin for 10 minutes at room temperature in 20mM HEPES pH7.5, 150 mM NaCl, 1 mM DTT, 1mM EDTA.

EM sample preparation

The intasome-DO2 NCP complex was formed and purified by size exclusion chromatography as previously described[15]. To allow strand transfer, the complex was incubated in the presence of 5 mM $MgCl_2$ for 30 minutes at room temperature. Cryo-EM sample preparation was performed as follows. 4 µl of the integration reaction were applied to plasma cleaned C-Flat 1/1 400 mesh grids. After 1 minute incubation, grids were double side blotted for 3.5 seconds 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 120kV, 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 into liquid nitrogen and later shipped into 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 µm. Data were collected as movies of 7 frames over 1.6 seconds giving a total applied dose of 56 electrons / $Å^2$. 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 minutes, prior to quenching with 50 mM TrisHCl pH 7.5. the complex was concentrated and buffer exchanged using a 50-kDa spin concentrator (Amicon) into 10 mM Tris-HCl pH 7, 20 mM NaCl, 1 mM EDTA, 1 mM DTT; 3.5 µl sample at 80 ng/ul (DNA concentration based on spectrophotometry) was added to Quantifoil 2/2 grids,

with fresh carbon evaporated onto the grids to better control ice thickness. Grids were glow discharged at 40mA for 1 minute. Sample was blotted in a Vitrobot Mark IV using -1 offset, 15 sec wait time and 2.5 sec blot at 4°C and 100% humidity, before plunge-freezing in liquid ethane. Grids were vitrified in liquid ethane and stored in liquid nitrogen prior to loading on a Titan Krios operated at 300kV. 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/Å^2 and a defocus range from -1.5 to -3.5 µm. 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 sec wait time and 3.5 sec blot at 4°C and 100% humidity, before plunge-freezing in liquid ethane. Data was acquired using a Falcon III detector operating in counting mode using a pixel size of 1.09 Å and total dose of 30 electrons / Å^2. A total of 1,300 Micrographs were collecting using automated EPU software.

Cryo-EM image processing

For the intasome-DO2 NCP complex dataset (**Supplementary Figure 1**), movie frames were corrected for to beam-induced drift[55] 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 Xmipp[56] and Relion version 1.3[57]. Contrast transfer function parameters were estimated using CTFFIND3[58], and all 2D and 3D classifications and 3D refinements were performed using RELION[57]. 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 map[15], filtered to 50 Å resolution, as a starting model. To speed up calculations, 8 classes were generated with a 15 degrees 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 degrees 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 previously[59] and the resulting map reached 3.9 Å resolution after correction for the modulation transfer function and sharpening[60]. Resolutions are reported according to the “gold-standard” Fourier Shell Correlation, using the 0.143 criterion[61].

For the D02-NCP-Streptavidin and Widom 601 NCP datasets (**Supplementary Figures 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 Gctf[62] 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 cryoSPARC[63]. 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-2.1[64] 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 (3UTB[65] from PDBredo) and PFV strand transfer complex (3OS0[16]) crystal structures were docked in the EM map using Chimera[66] 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_refine[67] with no additional geometry restraints. The geometry of resulting models was compared, and voxel/pixel size fine-tuned between 1.11 -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 Coot[68] and refined using phenix.real_space_refine (Nightly build version 1.10pre-2091)[69] and Namdinator[70,71]. Additional restraints describing protein secondary structure, DNA base pairing and stacking were used in Phenix. Protein geometry was assessed with Molprobity[67] and DNA geometry was assessed with 3DNA[72]. 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 density maps.

462

463 **Single-Molecule FRET experiments**

464 Doubly-labelled nucleosomes were generated with a biotin on distal exit DNA and a single
465 fluorophore donor (Cy3) attached on the proximal exit DNA end, and the acceptor
466 fluorophore (Cy5) at H2A position 119. To generate protein-Cy5-labelled octamers H2A
467 K119C was incorporated into octamers with H3.1 C96SC110A, H2B and H4 as described
468 above, with an additional desalting step in a Zeba Spin column (ThermoFisher, 7K MWCO)
469 to remove beta-mercaptoethanol. Octamers at 70 µM (140 µM of cysteine) were incubated
470 with 5 mM TCEP for 10 minutes at room temperature. To achieve partial labelling, sulpho-
471 Cy5 maleimide was added at 105 µM for 1 hour at room temperature. The reaction was
472 quenched by adding 5mM beta-mercaptoethanol and desalted to remove unreacted dye
473 (ThermoFisher, 7K MWCO). The extent of labelling was quantified by measuring the
474 595nm/280nm absorbance ratio, as well as by 2D intact mass ESI mass spectrometry. D02
475 DNA was generated by PCR, using oligos containing Biotin-TEG-C18 and Cy3 modifications
476 attached to the 5' termini. The PCR product was purified as described above. Nucleosomes
477 were reconstituted as described above.

478 Single-Molecule FRET experiments were performed with freshly purified intasome
479 complex. Quartz slides and coverslips were cleaned and passivated with methoxy-PEG-SVA
480 (M_r = 5,000, Laysan Bio, Inc.) containing 10% biotin-PEG-SVA (M_r = 5,000, Laysan Bio, Inc.)
481 in 100 mM sodium bicarbonate, and used to construct a microfluidic channel as described
482 previously[73]. Neutravidin (0.2 mg/ml in 50 mM Tris-HCl, pH 7.5, and 50 mM NaCl) was
483 injected in and incubated for 5 min. Excess neutravidin was washed out with intasome buffer
484 (25 mM bis-Tris propane, pH 7.45, 240 mM NaCl, 4 µM $ZnCl_2$ and 1 mM DTT). Biotinylated
485 fluorescently labelled nucleosomes in intasome buffer containing 0.2 mg/ml BSA were
486 surface immobilised by incubation in the microfluidic channel for 5 min. Excess nucleosomes
487 were washed out and immobilised nucleosomes imaged in imaging buffer composed of
488 intasome buffer in addition to 2 mM Trolox, oxygen scavenging system (2.5 mM 3,4-
489 dihydroxybenzoic acid, 250 nM protocatechuate dioxygenase) and 0.2 mg/ml BSA.
490 Experiments were performed in the absence and presence of 500 nM intasome and 5 mM
491 magnesium. Fluorescent molecules were imaged using a custom-built prism-based total-
492 internal reflection fluorescence (TIRF) microscope[74]. All measurements were recorded at
493 room temperature (21 ^{o}C) using continuous green laser (532 nm, 2.5 mW) excitation at 100
494 ms time resolution. Apparent FRET efficiencies were calculated as the ratio of acceptor
495 intensity divided by the sum of acceptor and donor intensities. FRET histograms of labelled
496 nucleosomes were obtained by calculating the mean FRET efficiency of 40-100 trajectories
497 from multiple fields of view, as stated in the figure captions. For each experimental condition,

sample was prepared twice. For each preparation, and for each condition, sample was measured several times over a period of two-three days.

Intasome strand-transfer and pull-down assays

Intasome integration assays were performed as described[15], briefly 5 µg of NCPs were incubated with 1.5 µg of intasome in intasome reaction buffer with and without 5mM $MgCl_2$ at 37°C for 15 minutes. 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 described[15].

Authors contributions

AC and PC initiated the study. DPM assembled the Intasome-NCP complex and LR determined the structure. DPM performed pull-down assays. MDW performed biochemistry, assembled NCP-D02-streptavidin and NCP-601 complexes and determined the structures. VEP built all models into the EM density maps. MG performed single molecule FRET experiments and data analysis, supervised by DSR. MDW, LR and AN performed cryo-EM imaging. MDW, PC and AC wrote the manuscript with input from the authors.

Acknowledgements

The authors woud like to thank Rishi Matadeen (formerly at NeCEN) for data collection of the NCP-intasome complex. The authors would like to 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). MDW 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).

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 6R0C 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 EMDB-4960 respectively.

Bibliography

1. Lesbats, P., Engelman, A.N. & Cherepanov, P. Retroviral DNA Integration. *Chem Rev* **116**, 12730-12757 (2016).
2. Li, M., Mizuuchi, M., Burke, T.R., Jr. & Craigie, R. Retroviral DNA integration: reaction pathway and critical intermediates. *EMBO J* **25**, 1295-304 (2006).
3. 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**, 6667-78 (1998).
4. Hare, S., Gupta, S.S., Valkov, E., Engelman, A. & Cherepanov, P. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **464**, 232-6 (2010).
5. Gerton, J.L., Herschlag, D. & Brown, P.O. Stereospecificity of reactions catalyzed by HIV-1 integrase. *J Biol Chem* **274**, 33480-7 (1999).
6. Hare, S., Maertens, G.N. & Cherepanov, P. 3'-processing and strand transfer catalysed by retroviral integrase in crystallo. *EMBO J* **31**, 3020-8 (2012).
7. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-60 (1997).
8. Richmond, T.J. & Davey, C.A. The structure of DNA in the nucleosome core. *Nature* **423**, 145-50 (2003).
9. Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D. & Klug, A. Structure of the nucleosome core particle at 7 A resolution. *Nature* **311**, 532-7 (1984).
10. Pryciak, P.M., Sil, A. & Varmus, H.E. Retroviral integration into minichromosomes in vitro. *EMBO J* **11**, 291-303 (1992).
11. Pryciak, P.M. & Varmus, H.E. Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell* **69**, 769-80 (1992).
12. 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 U S A* **91**, 5913-7 (1994).
13. 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**, 4704-14 (1994).
14. 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**, 1186-94 (2007).
15. Maskell, D.P. et al. Structural basis for retroviral integration into nucleosomes. *Nature* **523**, 366-9 (2015).
16. Maertens, G.N., Hare, S. & Cherepanov, P. The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature* **468**, 326-9 (2010).
17. Ballandras-Colas, A. et al. A supramolecular assembly mediates lentiviral DNA integration. *Science* **355**, 93-95 (2017).
18. Passos, D.O. et al. Cryo-EM structures and atomic model of the HIV-1 strand transfer complex intasome. *Science* **355**, 89-92 (2017).
19. Yin, Z. et al. Crystal structure of the Rous sarcoma virus intasome. *Nature* **530**, 362-6 (2016).
20. Ballandras-Colas, A. et al. Cryo-EM reveals a novel octameric integrase structure for betaretroviral intasome function. *Nature* **530**, 358-61 (2016).
21. Wigley, D.B. & Bowman, G.D. A glimpse into chromatin remodeling. *Nat Struct Mol Biol* **24**, 498-500 (2017).
22. Willhoft, O. et al. Structure and dynamics of the yeast SWR1-nucleosome complex. *Science* **362**(2018).
23. Cairns, B.R. Chromatin remodeling: insights and intrigue from single-molecule studies. *Nat Struct Mol Biol* **14**, 989-96 (2007).

24. Mackler, R.M. et al. Nucleosome DNA unwrapping does not affect prototype foamy virus integration efficiency or site selection. *PLoS One* **14**, e0212764 (2019).
25. Wilson, M.D. & Costa, A. Cryo-electron microscopy of chromatin biology. *Acta Crystallogr D Struct Biol* **73**, 541-548 (2017).
26. 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 a resolution. *J Mol Biol* **319**, 1097-113 (2002).
27. Bilokapic, S., Strauss, M. & Halic, M. Histone octamer rearranges to adapt to DNA unwrapping. *Nat Struct Mol Biol* **25**, 101-108 (2018).
28. 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).
29. Wilson, M.D. et al. The structural basis of modified nucleosome recognition by 53BP1. *Nature* **536**, 100-3 (2016).
30. Chua, E.Y. et al. 3.9 A structure of the nucleosome core particle determined by phase-plate cryo-EM. *Nucleic Acids Res* **44**, 8013-9 (2016).
31. 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**, 7822-32 (2011).
32. McGinty, R.K. & Tan, S. Nucleosome structure and function. *Chem Rev* **115**, 2255-73 (2015).
33. McGinty, R.K., Henrici, R.C. & Tan, S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**, 591-6 (2014).
34. Makde, R.D., England, J.R., Yennawar, H.P. & Tan, S. Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature* **467**, 562-6 (2010).
35. Arnaudo, N. et al. The N-terminal acetylation of Sir3 stabilizes its binding to the nucleosome core particle. *Nat Struct Mol Biol* **20**, 1119-21 (2013).
36. McGinty, R.K. & Tan, S. Recognition of the nucleosome by chromatin factors and enzymes. *Curr Opin Struct Biol* **37**, 54-61 (2016).
37. Zhou, K., Gaullier, G. & Luger, K. Nucleosome structure and dynamics are coming of age. *Nat Struct Mol Biol* (2018).
38. Anderson, C.J. et al. Structural Basis for Recognition of Ubiquitylated Nucleosome by Dot1L Methyltransferase. *Cell Rep* **26**, 1681-1690 e5 (2019).
39. 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).
40. Valencia-Sanchez, M.I. et al. Structural Basis of Dot1L Stimulation by Histone H2B Lysine 120 Ubiquitination. *Mol Cell* (2019).
41. 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* (2019).
42. Bilokapic, S., Strauss, M. & Halic, M. Structural rearrangements of the histone octamer translocate DNA. *Nat Commun* **9**, 1330 (2018).
43. Kujirai, T. et al. Structural basis of the nucleosome transition during RNA polymerase II passage. *Science* **362**, 595-598 (2018).
44. Ehara, H. et al. Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science* **363**, 744-747 (2019).
45. 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).
46. Farnung, L., Vos, S.M., Wigge, C. & Cramer, P. Nucleosome-Chd1 structure and implications for chromatin remodelling. *Nature* **550**, 539-542 (2017).
47. Eustermann, S. et al. Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* **556**, 386-390 (2018).
48. Ayala, R. et al. Structure and regulation of the human INO80-nucleosome complex. *Nature* **556**, 391-395 (2018).

49. Sundaramoorthy, R. et al. Structure of the chromatin remodelling enzyme Chd1 bound to a ubiquitinylated nucleosome. *Elife* **7**(2018).
50. Sundaramoorthy, R. et al. Structural reorganization of the chromatin remodeling enzyme Chd1 upon engagement with nucleosomes. *Elife* **6**(2017).
51. 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* (2019).
52. Li, M. et al. Mechanism of DNA translocation underlying chromatin remodelling by Snf2. *Nature* **567**, 409-413 (2019).
53. Lesbats, P. et al. Functional coupling between HIV-1 integrase and the SWI/SNF chromatin remodeling complex for efficient in vitro integration into stable nucleosomes. *PLoS Pathog* **7**, e1001280 (2011).
54. Dyer, P.N. et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol* **375**, 23-44 (2004).
55. Li, X. et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* **10**, 584-90 (2013).
56. 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).
57. Scheres, S.H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-30 (2012).
58. Mindell, J.A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* **142**, 334-47 (2003).
59. 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).
60. Rosenthal, P.B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* **333**, 721-45 (2003).
61. Scheres, S.H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat Methods* **9**, 853-4 (2012).
62. Zhang, K. Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12 (2016).
63. 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).
64. Kimanius, D., Forsberg, B.O., Scheres, S.H. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *Elife* **5**(2016).
65. 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**, 6338-52 (2012).
66. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).
67. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).
68. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
69. Afonine, P.V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544 (2018).
70. 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**, 174-80 (2009).
71. Kidmose, R.T. et al. Namdinator - Automatic Molecular Dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *bioRxiv 501197*.

703 72. Lu, X.J. & Olson, W.K. 3DNA: a software package for the analysis, rebuilding and
704 visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* **31**,
705 5108-21 (2003).
706 73. Lamichhane, R., Solem, A., Black, W. & Rueda, D. Single-molecule FRET of protein-
707 nucleic acid and protein-protein complexes: surface passivation and immobilization.
708 *Methods* **52**, 192-200 (2010).
709 74. Zhao, R. & Rueda, D. RNA folding dynamics by single-molecule fluorescence
710 resonance energy transfer. *Methods* **49**, 112-7 (2009).
711