1

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

75   repositioning of the nucleosomal DNA along the histone octamer. In addition, a systematic
76   analysis is needed to understand potential role of histone tails in intasome engagement.

77   Herein, we employ a combination of cryo-EM and single-molecule FRET (smFRET) to
78   understand what impact retroviral integration has on the structure of the target NCP. We find
79   that strand transfer causes both nucleosomal DNA looping as well as sliding by two base
80   pairs along the histone octamer. With our new findings we uncover unexpected similarities
81   between the mechanisms of retroviral integration and ATP-dependent chromatin remodeling
82   [21-23].

83

Results

85

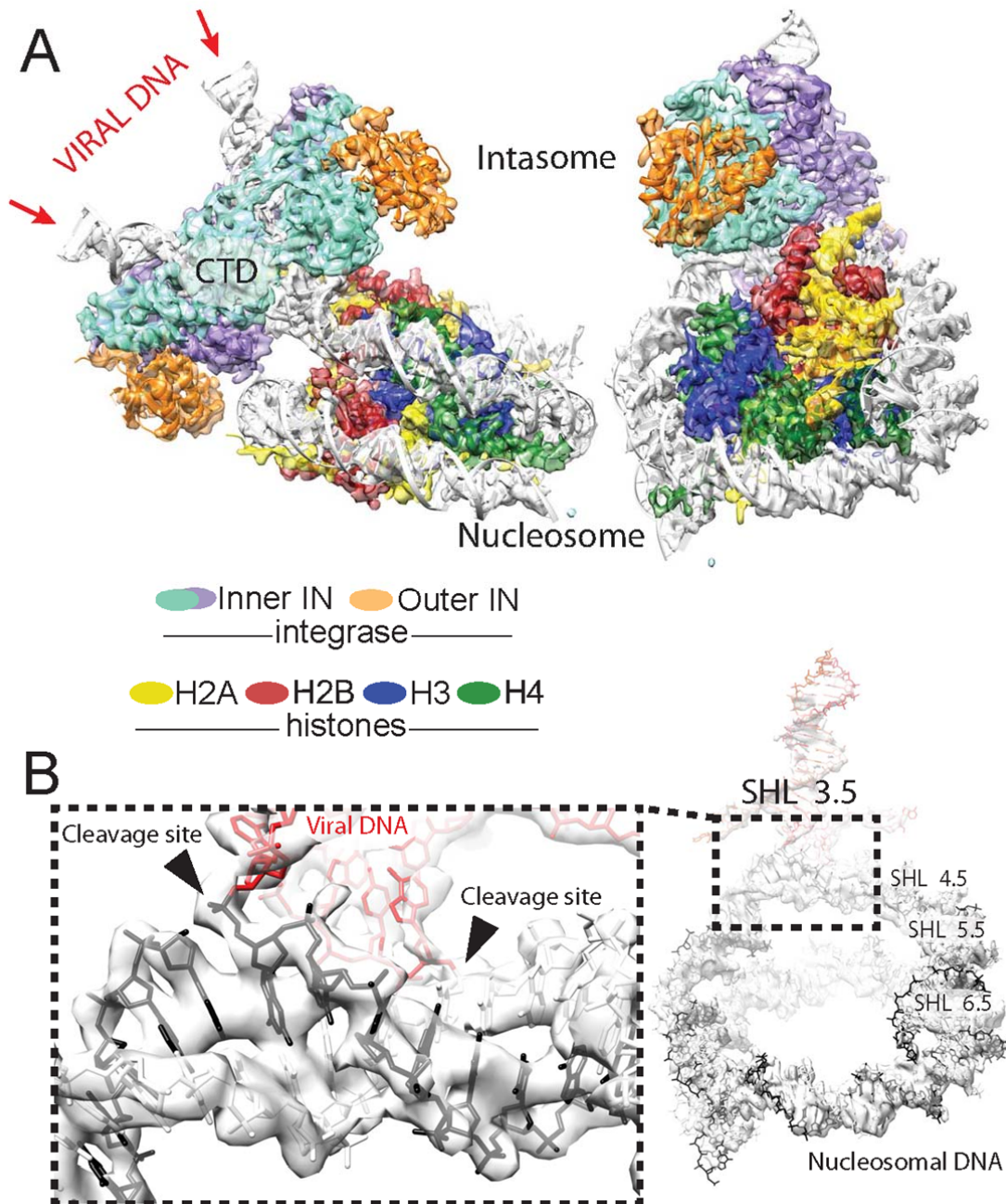**Structure of Intasome-NCP strand-transfer complex**

87   To understand intasome strand transfer into NCPs, we assembled the complex of the PFV
88   intasome and the NCP containing a native human DNA sequence (termed D02), selected for
89   its ability to form a stable PFV-NCP complex[15]. Following isolation by size exclusion
90   chromatography, the intasome-NCP complex was incubated in the presence of Mg^{2+} to
91   facilitate strand transfer[15]. We then used cryo-EM imaging and single-particle approaches to
92   determine the structure of the resulting post-catalytic assembly to 3.9 Å resolution
93   (**Supplementary Figure 1** and **Supplementary Table 1**). Docking known crystallographic
94   coordinates into the cryo-EM map and use of real-space refinement allowed us to generate
95   an atomic model of the Intasome-NCP strand transfer complex.

96   As previously observed, intasome engages the strongly preferred site on the nucleosomal
97   DNA, at SHL 3.5 [15,24] (**Figure 1**). The new structure is overall similar to the original lower-
98   resolution intasome-NCP complex, which was captured in the pre-catalytic state (**Figure
99   1A**), confirming that strand transfer is not accompanied by large conformational
100 rearrangements[6]. According to the atomic model, at the integration site, DNA is lifted by 7 Å
101 from the histone octamer and bent to allow access to the IN catalytic centre, in excellent
102 agreement with the earlier observations based on the crystal structure of the PFV strand
103 transfer complex and the lower-resolution intasome-NCP cryo-EM data.

104 Local resolution is highest (~3.5 Å) for the histone octamer core, while DNA density on the
105 outer perimeter tends to decrease (~4-4.5 Å), as observed previously for other NCP
106 structures (**Supplementary Figure 1**) [25-27]. Nevertheless, we could confidently model the
107 DNA phosphate backbone for the entire assembly. The integration site on the nucleosomal
108 DNA is sandwiched between the histones and the intasome, resulting in higher local
109 resolution (~3.7 Å). Notably, a discontinuity in the cryo-EM density resulting from the

110 nucleosomal DNA cleavage at the site of integration (**Figure 1B**) confirms that strand

111 transfer has indeed occurred in our nucleoprotein assembly as observed biochemically[15]

112 (**Supplementary Figure 2**).

113

115 **Figure 1** Intasome-NCP strand-transfer complex visualised by cryo-EM. **(A)** 3.9 Å resolution structure

116 of the post-catalytic intasome-NCP complex. **(B)** Covalently linked viral (red) and nucleosomal (black)

117 DNA. Integration occurs at SHL location 3.5.

118 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

124   and one outer IN subunit (**Figure 1A**). Near the integration site, the alpha C-helix of histone

125   H2B makes direct contact with the C-terminal domain of one catalytically competent IN

126   subunit, providing corroborating evidence for the reported role of IN residues Pro135,

127   Pro239 and Thr240 in engaging C-terminal H2B [15]. The higher quality of the new cryo-EM

128   map allowed us to build a backbone model for a segment of the N-terminal H2A tail,

129   revealing close proximity between positively charged Lys-9 and Arg-11 on H2A and the

130   catalytically competent IN C-terminal domain (**Figure 2A**). Coherently, we observe that

131   complex formation is reduced with an NCP containing a histone H2A truncation of the first

132   12, but not the first 8 residues (**Figure 2B**). Furthermore, Ala substitutions of either H2A at

133   Lys-9 or Arg-11 affect complex stability, while a combination of the two substitutions fully

134   abrogates the stable complex formation under conditions of the pull-down assay **(Figure**

135   **2C)**.
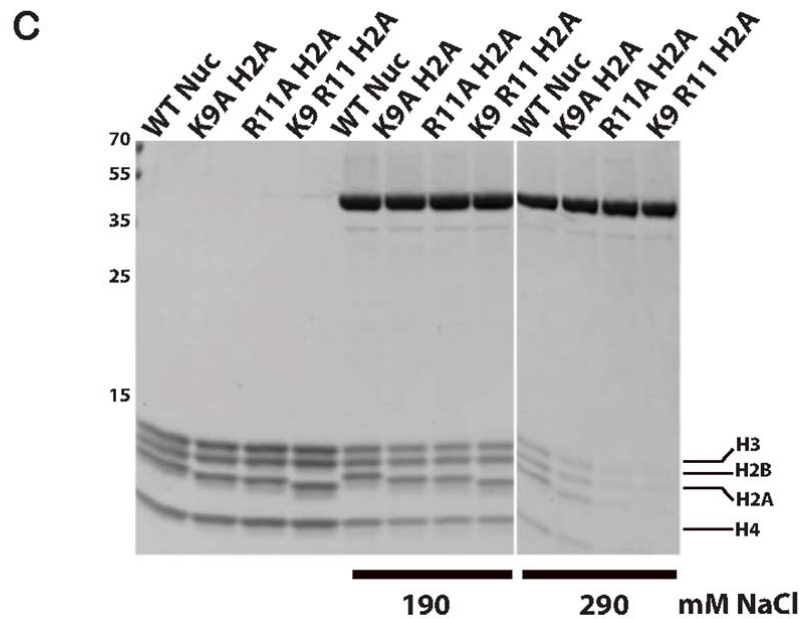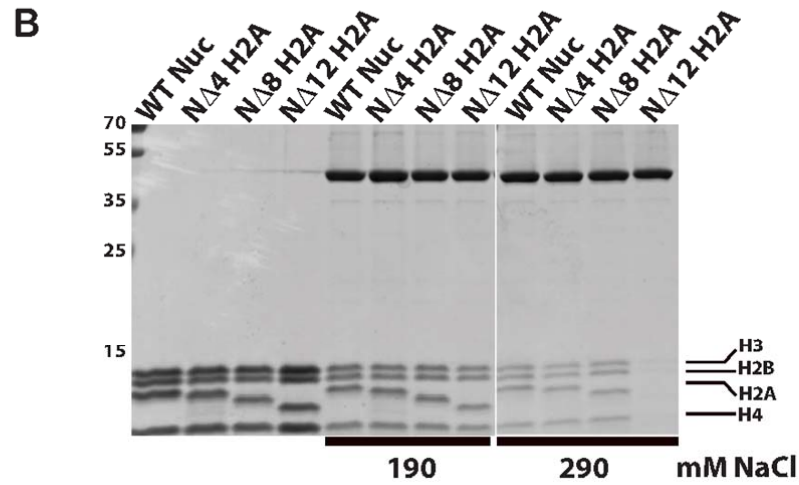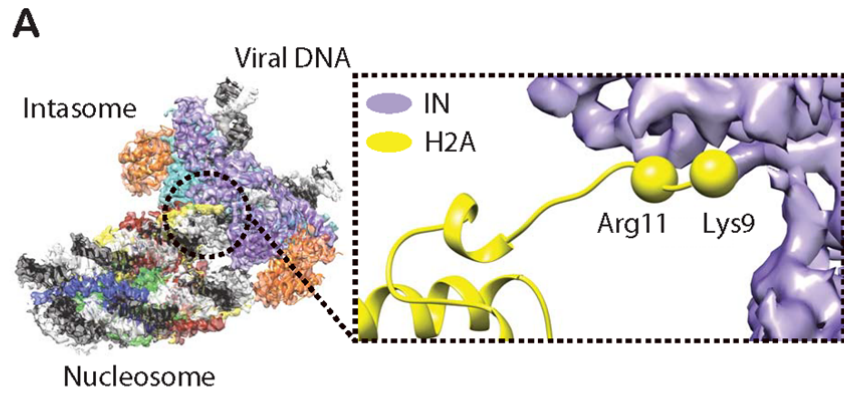
Figure 2 Intasome interaction with the N-terminal histone H2A tail. (**A**) Histone H2A residues Lys-9 and Arg-11 play a key role in intasome-NCP interaction. (**B**) Pull-down assay with immobilised intasome binding to NCP H2A tail deletion variants. (**C**) Pull-down assays with immobilised intasome binding to NCP H2A N-terminal tail variants containing single- or double-point mutations.

142   **Asymmetric reconstruction of isolated NCP bearing the human D02 DNA sequence**

143   Similar to the pre-catalytic complex, our new structure of an intasome-NCP strand-transfer

144   complex features a nucleosomal DNA loop bulging away from the protein octamer by ~7 Å at

145   the integration site. Although occurring at a different superhelical location, the DNA looping

146   is reminiscent of structures of NCPs engaged by chromatin remodelers such as SWR1.

147   Interestingly, DNA looping by SWR1 is accompanied by both sliding of nucleosomal DNA, as

148   well as histone octamer distortion [22]. We wanted to test whether intasome-induced looping is

149   compensated by nucleosomal DNA sliding along the histone octamer, as observed for

150   chromatin remodelers. To this end, we decided to directly compare and contrast the cryo-EM

151   structure of the intasome-NCP strand-transfer complex with that of an isolated NCP,

152   containing the same native human D02 nucleosomal DNA sequence[15].

153   Reconstructing a D02 NCP presented a number of significant challenges. Firstly, the NCP

154   containing D02 DNA is less stable than NCPs wrapped with strongly positioning sequences

155   such as Widom 601[15,28]. Our EM analysis of the isolated NCP D02 revealed that, unlike the

156   intasome complex, D02 NCPs had the tendency to become unravelled, especially in the

157   presence of higher salt (data not shown). However, exposure to mild crosslinking conditions

158   (0.05% glutaraldehyde, 5 min, 4°C) yielded tractable particles that were visible on open-hole

159   cryo grids. Importantly, mild NCP-crosslinking did not prevent intasome activity as measured

160   in strand-transfer assays (**Supplementary Figure 2**). A second challenge was presented by

161   the asymmetry of the D02 DNA sequence, which leads to the strongly preferred intasome

162   capture at one side of the NCP[15]. Thus, to describe any intasome-dependent sliding along

163   the histone octamer, we first had to reconstruct the D02 NCP avoiding two-fold averaging.

164   However, both the histone octamer and the DNA backbone contain a prominent two-fold

165   symmetric character, which strongly influence particle alignment and prevent asymmetric

166   reconstruction. To facilitate asymmetric particle alignment, we introduced a biotin moiety on

167   the end of the DNA arm distal from the integration site and decorated NCPs with streptavidin

168   (**Figure 3A**). Critically, streptavidin attachment did not affect NCP stability, nor the ability of

169   intasome to integrate into NCPs (**Supplementary Figure 2**). Crosslinked D02 NCPs,

170   imaged by cryo-EM and analysed by two-dimensional (2D) averaging, revealed multiple

171   views of the coin-shaped NCP assemblies (**Figure 3B**). Particles appeared decorated by

172   diffuse density projecting from one DNA arm, which we assigned to streptavidin. Free

173   streptavidin particles (~75 kDa) could also be identified amongst the 2D class averages

174   (**Supplementary Figure 3**). Next, we used single-particle reconstruction to determine the

175   4.2 Å resolution structure of NCP-D02-streptavidin complex (**Supplementary Figure 3** and

176   **4**). As the streptavidin is linked to the 5' end of a distal DNA arm, it is less ordered than the

177   rest of the assembly, and appears not to be engaged in any stabilising interaction with the

178
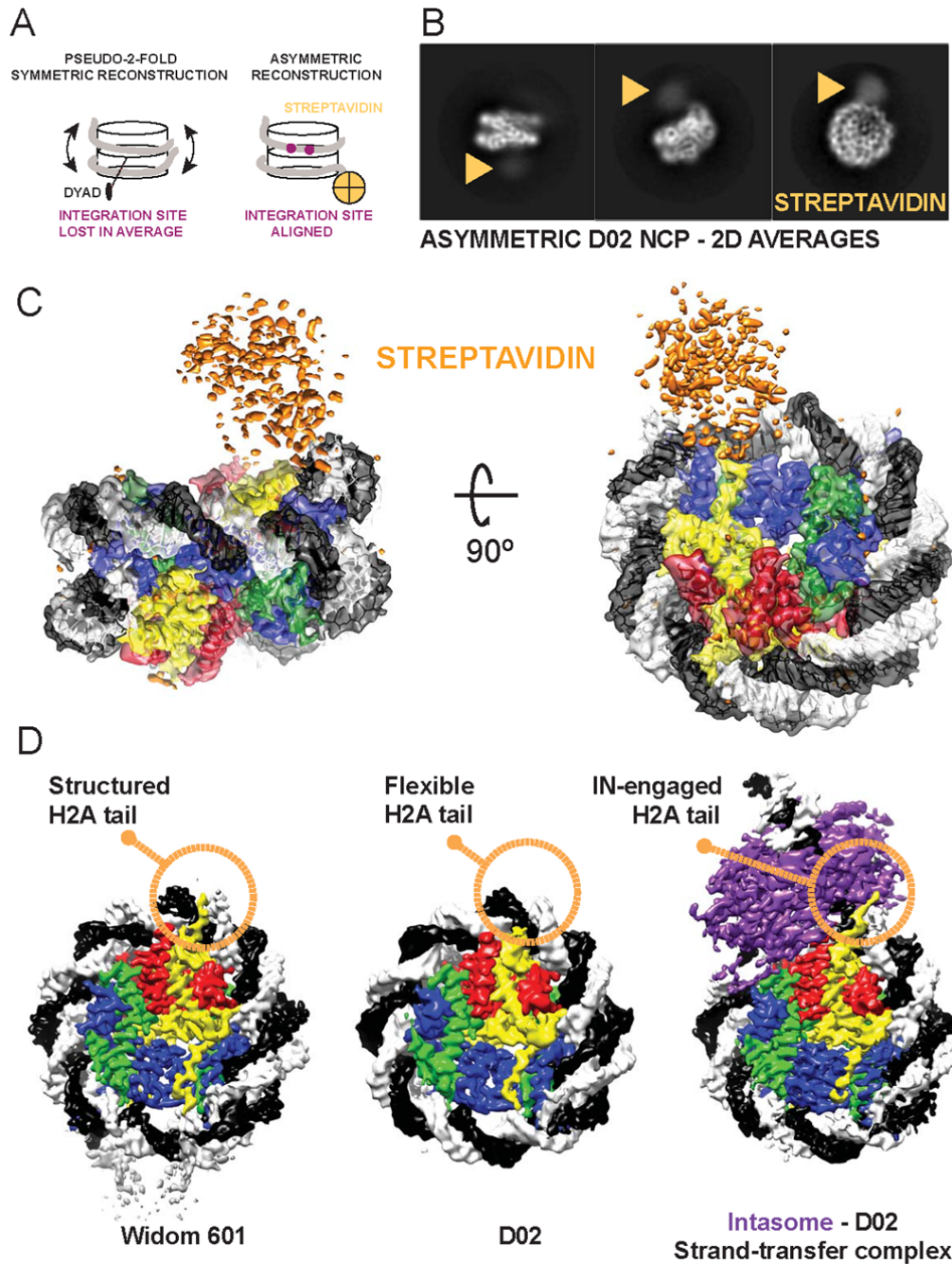
Figure 3 Asymmetric reconstruction of the isolated D02 NCP. **(A)** Streptavidin labelling allows asymmetric reconstruction of the D02 NCP, which avoids pseudo-two-fold symmetry averaging of the integration-site DNA. **(B)** 2D averages of labelled D02 NCP particles reveal a discernible, diffuse density 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).

188 NCP core (**Figure 3C** and **Supplementary Figure 3**). Therefore, streptavidin helps align
189 particles asymmetrically while seemingly not interfering with the NCP structure.

190 Originally selected from a genome-wide screen for strong intasome interactors, the D02
191 DNA sequence allowed isolation of a mono-disperse intasome-NCP complex[15]. Detailed
192 inspection of the isolated D02 NCP cryo-EM maps provides insight into intasome selectivity.
193 Firstly, nucleosomal DNA arms appear to be flexible (as detected by inspection of the local
194 resolution map reported in **Supplementary Figure 3**, and given the significant number of
195 unwrapped NCPs averaged during analysis). We asked whether the same flexibility could be
196 observed for a NCP containing a strong positioning sequence such as Widom 601, which is
197 a poor substrate for intasome binding and integration. To this end, we solved the 3.5 Å
198 resolution cryo-EM structure of a Widom-601-wrapped nucleosome containing strongly
199 positioned Widom-601 sequence with 13-bp long linker DNA arms (**Supplementary Figure
200 5**). Only linker DNA fragments display a degree of flexibility in the Widom 601 structure. We
201 postulated at this stage that flexible NCP arms in D02 might favour nucleosomal DNA
202 repositioning required to support DNA looping, prompting us to further investigate the
203 mechanism.

204 A second notable feature in the D02 NCP structure is the limited interaction between DNA
205 and the N-terminal tail of H2A, reflected by poorly defined density contacting nucleosomal
206 DNA at SHL 4.5. This differs for example from our structure of Widom 601 NCP, which
207 shows discrete ordering of H2A N-terminal tail in the minor groove of nucleosomal DNA at
208 the equivalent position, in agreement with previous crystallography and cryo-EM studies [7,29-
209 31]. We speculate that loose DNA-engagement renders the histone H2A tail available for
210 intasome binding as observed in our strand-transfer complex, hence improving substrate
211 selection (**Figure 3D**).

212 **Retroviral integration causes a shift in nucleosomal DNA register**

213 To understand the impact of retroviral integration on NCP architecture, we analysed the
214 structural changes in the NCP that accompany productive engagement with the intasome.
215 Comparison of the intasome-D02 NCP structures prior to and after strand transfer shows
216 that histones undergo relatively minor distortions with a global r.m.s. deviation of backbone
217 atom positions of 2.7 Å (**Supplementary Figure 6**). Conversely, in our atomic model DNA
218 looping at the integration site is compensated by a significant change in nucleosomal DNA
219 register, with the nucleosomal DNA arm proximal to the integration site shifting by 2 bp
220 (**Figure 4A**). This shift in register extends from SHL 7 to SHL 2.5, where an interaction with
221 H3 element L1 appears to hold DNA in place and limit downstream sliding of the double
222 helix (**Figure 4B** and **Supplementary Movie 1**).
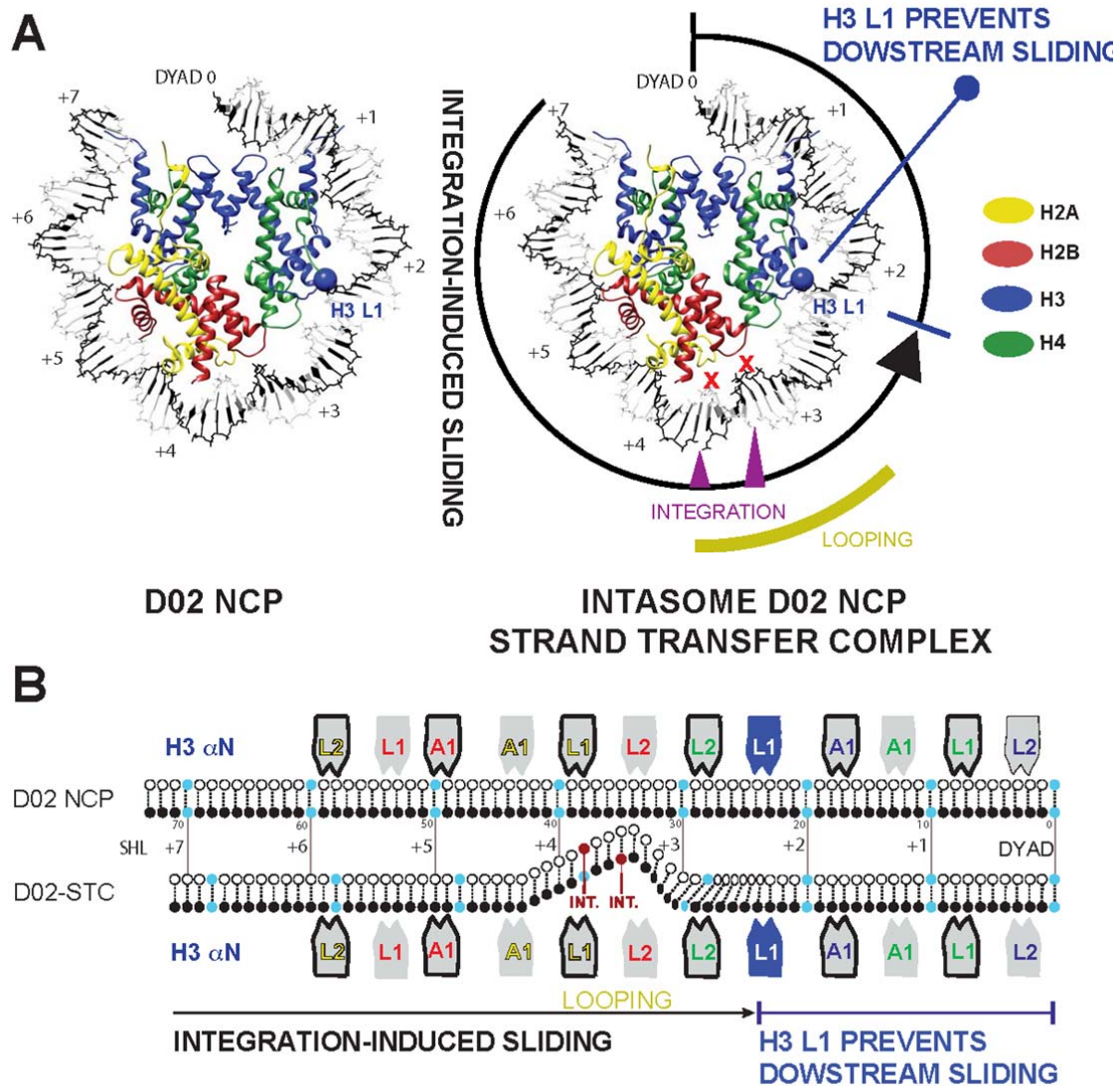
Figure 4 Integration-promoted DNA sliding observed by cryo-EM. **(A)** On the left, slabbed view of the isolated D02 NCP. Histone H3 L1 element is highlighted with a blue ball. On the right, DNA looping required for retroviral integration causes a shift in the DNA register, which extends from SHL 7 to 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.

263
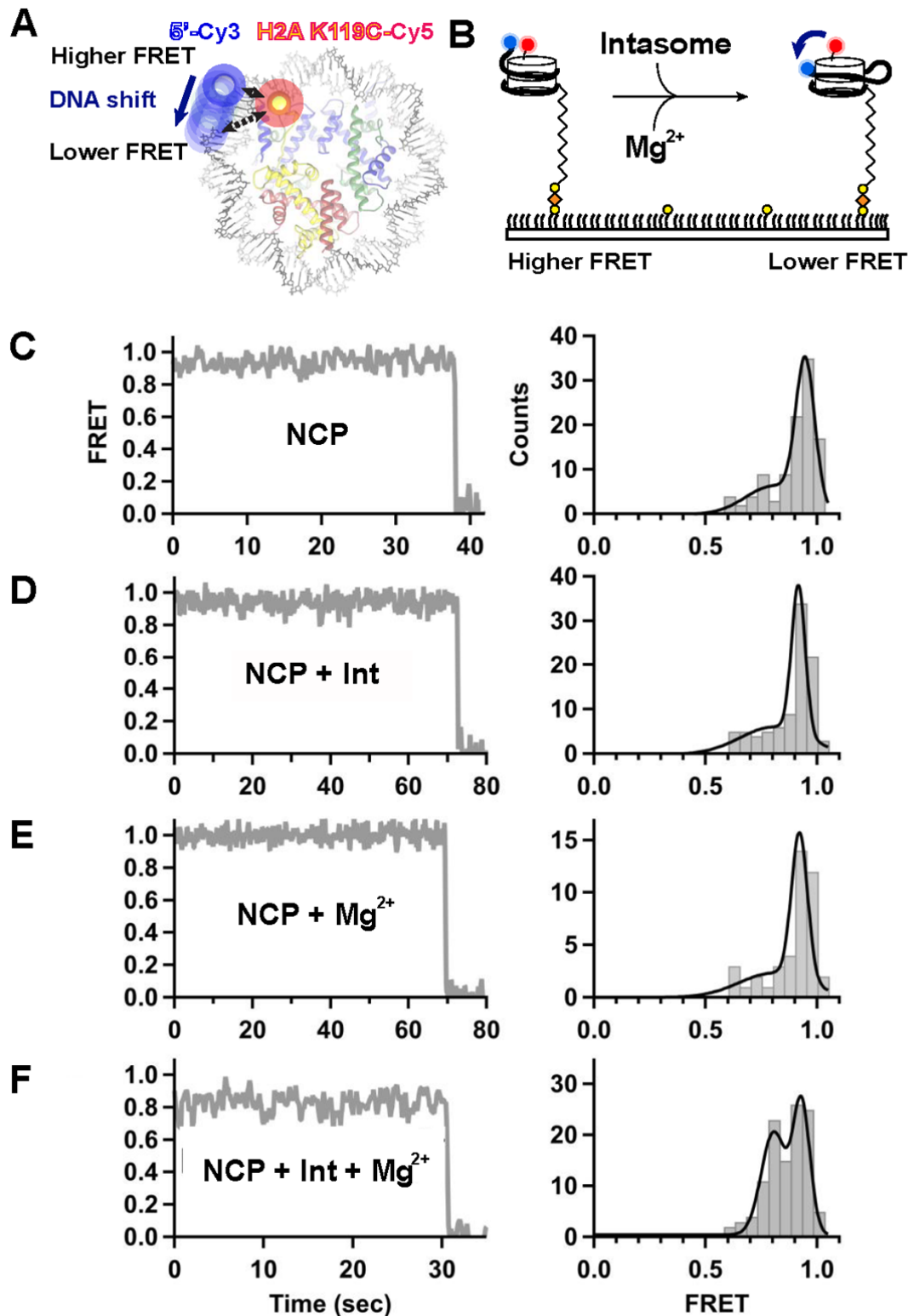
Figure 5 Integration-promoted nucleosomal DNA sliding observed by single molecule FRET. (**A**) Fluorescently labelled NCP: labelled octamer (**Cy5** on H2A, red) wrapped with biotinylated labelled DNA (**Cy3** on exit site, blue). (**B**) NCPs are surface-immobilised on neutravidin (orange) coated biotin-PEG (yellow) slides. Intasome-induced translocation in the presence of magnesium is detected as a

268   FRET decrease. Representative single-molecule FRET trajectories (left) and histograms (right) of
269   NCPs (**C**) without intasome or magnesium (N = 105), (**D**) in the presence of intasome (N = 93), (**E**) in
270   the presence of magnesium (N = 42), and (**F**) in presence of both intasome and magnesium (N =
271   115). A significant population shift to ~0.8 FRET in the presence of both intasome and magnesium.
272   Data were collected at 100 ms/frame (10 Hz) and smoothed with a 3-point moving average. Black
273   lines in FRET histogram are fits to two Gaussian distributions.

274   **Discussion**

275   Over the last 35 years macromolecular crystallography has provided several high-resolution
276   views of the NCP and its binding partners. These efforts led to describing the NCP
277   architecture at an atomic level[7-9], explained how DNA sequence can influence wrapping of
278   the double helix[32], and how common docking sites on the histone octamer are recognised by
279   different interactors[33-37]. Over the last four years, cryo-EM has started to provide a dynamic
280   view of the NCP [25,38-41]. Recent data indicated that NCPs are more flexible in solution, with
281   the histone octamer visiting more compacted or extended states, compared with a
282   nucleosome trapped in a crystal lattice[42]. NCP unwrapping has been visualised with cryo-
283   EM, for example in the context of the hexasome, which is an NCP with partially unpeeled
284   DNA, due to the loss of one H2A/H2B dimer[27]. Spectacular views of progressively
285   unwrapped NCPs have been obtained for transcribing RNA polymerase II captured during
286   NCP passage [43,44]. Moreover, cryo-EM provided the first glimpses of ATP-dependent NCP
287   translocation through a mechanism involving DNA looping and sliding along the histone
288   octamer [22,38,45-51].

289   Our high-resolution view of a post-catalytic intasome-NCP complex provides a new example
290   of a local remodeling of nucleosomal DNA. Although previous work established formation of
291   a DNA loop during productive intasome-NCP interaction, it was not clear whether the loop
292   forms by local stretching of the double helix or through a shift in nucleosomal DNA register.
293   Because IN must catalyse only one strand transfer event and does not need to cycle
294   between states on the chromatin, it does not depend on a power source, unlike ATP-driven
295   translocases and NCP remodelers. Therefore, all conformational DNA rearrangements are
296   offset by energy released with the formation of the intasome-NCP interface. Nevertheless,
297   similarities with the mechanism of DNA translocation of chromatin remodelers can be
298   identified. In fact, in both systems, DNA is looped out of the histone core, causing a
299   compensatory register shift of the double helix wrapped around the octamer. Nucleosomal
300   DNA looping at SHL 3.5 is required for access to the IN active site[15], and causes DNA
301   sliding around the histone octamer, with global repositioning extending from SHL 7 to SHL 2.
302   At this site, histone H3 element L1 holds the sugar-phosphate backbone in place, preventing
303   any further downstream shift in DNA register (**Figure 4B** and **Supplementary Movie 1**).

304   Using cryo-EM, Kurumizaka and colleagues have recently shown that the same H3 L1-DNA

305   interaction stalls RNA polymerase II during NCP passage[43]. ATP-powered translocases

306   such as Swr1 and Snf2 have been observed to engage and loop out SHL 2 DNA, disrupting

307   the H3 L1-DNA interaction [22,45,52]. It is tempting to speculate that the concerted action of IN

308   and SHL 2 remodelers could act synergistically during DNA unpeeling and strand-transfer

309   complex disassembly, required to complete retroviral integration. In support of this notion, a

310   functional coupling during retroviral integration has been reported between HIV-1 integrase

311   and Snf2-related SWI/SNF chromatin remodelling complex [53].

312

313

314   **Supplementary Movie 1:** Nucleosome core particle morphed between the isolated and the strand

315   transfer complex state.

316
317

318 **Materials and Methods**

319

320 **Intasome purification**

321 The intasome was assembled using recombinant PFV IN and double stranded synthetic

322 oligonucleotides mimicking the pre-processed U5 end of the vDNA as previously

323 described[4,15]. Briefly, hexahistidine-tagged IN was overexpressed in BL-21 CodonPlus RIL

324 cells (Agilent). Cells were lysed in 25 mM Tris-HCl pH 7.4, 0.5 M NaCl, 1 mM PMSF by

325 sonication; clarified lysate supplemented with 20 mM imidazole was applied to packed,

326 equilibrated Ni-NTA resin (Qiagen). The resin and washed extensively in lysis buffer

327 supplemented with 20 mM imidazole. Bound proteins were eluted with lysis buffer

328 supplemented with 200 mM imidazole and protein-containing fractions were supplemented

329 with 5 mM DTT. The hexahistidine-tag was cleaved by incubation with human rhinovirus 14

330 3C protease. The protein, diluted to reduce the NaCl concentration to 200 mM, was loaded

331 onto a HiTrap Heparin column (GE Healthcare). IN was eluted using a linear gradient of

332 0.25-1 M NaCl. IN-containing fractions were concentrated and further purified by size

333 exclusion chromatography through a Superdex-200 column (GE Healthcare), equilibrated in

334 25 mM Tris pH 7.4, 0.5 M NaCl. Protein, supplemented with 10% glycerol and 10 mM DTT,

335 was concentrated to 10 mg/ml, as estimated by spectrophotometry at 280 nm and stored at -

336 80°C.

337 To assemble the intasome a mixture containing 120 μM PFV IN and 20 μM pre-annealed

338 DNA oligonucleotides 5'-TGCGAAATTCCATGACA and 5'-ATTGTCATGGAATTTCGCA

339 (IDT) in 500 mM NaCl was dualized against 50mM BisTris propane-HCl pH 7.45, 200 mM

340 NaCl, 40 μM ZnCl$_2$, 2 mM DTT for 16 h at 18°C. Following dialysis, the assembly reaction,

341 supplemented with NaCl to a final concentration of 320 mM, was incubated on ice for 1 h

342 prior to purification on Superdex-200 column in 25 mM Bis-Tris propane-HCl pH 7.45, 320

343 mM NaCl. Purified intasome, concentrated by ultrafiltration, was kept on ice for immediate

344 use.

345

346 **NCP formation**

347 NCPs were assembled essentially as described [15,54]. Briefly Human H2A, H2A K119C, H2B,

348 H3.3, H3.1 C96SC110A and H4 were over-expressed in *E.coli* and purified from inclusion

349 bodies. Histones were refolded from denaturing buffer through dialysis against 10 mM Tris-

350 HCl pH 7.5, 2 M NaCl, 5 mM beta-mercaptoethanol, 1 mM EDTA buffer, and octamers were

351 purified by size exclusion chromatography over a Superdex-200 column (GE Healthcare).

352 DNA fragments for wrapping NCPs (171-bp Widom-601 DNA, 145-bp D02 DNA or D02 DNA

353 appended with biotin and fluorophores) were generated by PCR using Pfu polymerase and

354   HPLC-grade oligonucleotides (IDT). PCR products generated in 96-well plates (384 x 100 µl)

355   were pooled, filtered and purified on a ResorceQ column as described [15]. NCPs were

356   assembled by salt dialysis as described [15,29,54] and heat repositioned at 37°C for 30 minutes.

357   D02 containing NCPs were further purified using a PrepCell apparatus with a 5%

358   polyacrylamide gel (BioRad).

359

NCP-streptavidin complex formation

361   Purified *Streptomyces avindii* streptavidin powder (Sigma-Aldrich) was resuspended in 20

362   mM HEPES-NaOH pH 7.5, 150 mM NaCl at a final concentration of 35 µM. A derivative of

363   D02 DNA was used for NCP reconstitution, containing a 5' biotin moiety on the exit arm

364   distal from the intasome-engagement site. To form the NCP-streptavidin complex,

365   biotinylated D02 NCP (0.5 µM) was incubated with 0.3 µM streptavidin for 10 minutes at

366   room temperature in 20mM HEPES pH7.5, 150 mM NaCl, 1 mM DTT, 1mM EDTA.

367

EM sample preparation

369   The intasome-DO2 NCP complex was formed and purified by size exclusion

370   chromatography as previously described[15]. To allow strand transfer, the complex was

371   incubated in the presence of 5 mM $MgCl_2$ for 30 minutes at room temperature. Cryo-EM

372   sample preparation was performed as follows. 4 µl of the integration reaction were applied to

373   plasma cleaned C-Flat 1/1 400 mesh grids. After 1 minute incubation, grids were double side

374   blotted for 3.5 seconds using a CP3 cryo-plunger (Gatan), operated at 80% humidity, and

375   quickly plunge-frozen into liquid ethane. Ice quality was checked using a JEOL-2100 Lab6

376   operated at 120kV, using a 914 side-entry cryo-holder (Gatan), and images were recorded

377   on an UltraScan 4kx4k camera (Gatan). The best cryo-grids were retrieved, stored into liquid

378   nitrogen and later shipped into a dry-shipper to NeCEN (University of Leiden, The

379   Netherlands). At NeCEN, grids were loaded into a Cs corrected Titan Krios microscope and

380   the data was collected over two different sessions using the EPU software (ThermoFisher

381   Scientific). Images were recorded at a nominal magnification of 59,000 X on Falcon II direct

382   electron detector yielding a pixel size of 1.12 Å / pixel with a defocus range of -1.5 to -3.5

383   µm. Data were collected as movies of 7 frames over 1.6 seconds giving a total applied dose

384   of 56 electrons / $Å^2$. A total of 4,916 movies were collected.

385   The D02 NCP biotin-streptavidin complex was gently cross-linked with 0.05% glutaldehyde

386   at room temperature for 5 minutes, prior to quenching with 50 mM TrisHCl pH 7.5. the

387   complex was concentrated and buffer exchanged using a 50-kDa spin concentrator

388   (Amicon) into 10 mM Tris-HCl pH 7, 20 mM NaCl, 1 mM EDTA, 1 mM DTT; 3.5 µl sample at

389   80 ng/ul (DNA concentration based on spectrophotometry) was added to Quantifoil 2/2 grids,

390   with fresh carbon evaporated onto the grids to better control ice thickness. Grids were glow

391   discharged at 40mA for 1 minute. Sample was blotted in a Vitrobot Mark IV using -1 offset,

392   15 sec wait time and 2.5 sec blot at 4°C and 100% humidity, before plunge-freezing in liquid

393   ethane. Grids were vitrified in liquid ethane and stored in liquid nitrogen prior to loading on a

394   Titan Krios operated at 300kV. Data was acquired using a Falcon III detector operating in

395   counting mode using a pixel size of 1.09 Å, a total dose of 30 electrons/Å^2 and a defocus

396   range from -1.5 to -3.5 µm. A total of 4,182 movies were collected automatically using the

397   EPU software (ThermoFisher Scientific). The Widom 601 NCP sample was applied to freshly

398   glow discharged Quantifoil 2/2 grids and sample was blotted in a Vitrobot Mark IV using -1

399   offset, 10 sec wait time and 3.5 sec blot at 4°C and 100% humidity, before plunge-freezing

400   in liquid ethane. Data was acquired using a Falcon III detector operating in counting mode

401   using a pixel size of 1.09 Å and total dose of 30 electrons / Å^2. A total of 1,300 Micrographs

402   were collecting using automated EPU software.

403

404   **Cryo-EM image processing**

405   For the intasome-DO2 NCP complex dataset (**Supplementary Figure 1**), movie frames

406   were corrected for to beam-induced drift[55] and a sum of each aligned movie was used in the

407   first steps of image processing. All movies showing any remaining drift or containing ice

408   were discarded at this stage, and only the best 3,125 movies were selected for further image

409   processing. First, 989,177 particles were automatically picked using Xmipp[56] and Relion

410   version 1.3[57]. Contrast transfer function parameters were estimated using CTFFIND3[58], and

411   all 2D and 3D classifications and 3D refinements were performed using RELION[57]. After 2

412   rounds of 25 iterations of 2D classification, 335,989 particles remained and were subjected

413   to 3D classification using the pre-catalytic intasome-NCP map[15], filtered to 50 Å resolution,

414   as a starting model. To speed up calculations, 8 classes were generated with a 15 degrees

415   angular sampling. The best 3 classes were merged into one 232,000 particles dataset. 3d

416   refinement of this subset yielded a 4.7 Å map. A second round of 3D classification step was

417   performed with 4 classes and a finer 7.5 degrees angular sampling. The best 3 classes were

418   merged together for a total of 177,155 particles. Refinement of this dataset yielded a 4.2 Å

419   map. Statistical movie processing was then performed as described previously[59] and the

420   resulting map reached 3.9 Å resolution after correction for the modulation transfer function

421   and sharpening[60]. Resolutions are reported according to the "gold-standard" Fourier Shell

422   Correlation, using the 0.143 criterion[61].

423   For the D02-NCP-Streptavidin and Widom 601 NCP datasets (**Supplementary Figures 3-5**)

424   all micrographs were motion-corrected using MotionCorr2 using all frames (D02-NCP-

425   Streptavidin) or removing the first frame (Widom 601 NCP). CTF parameters were estimated

426   using Gctf[62] and poor micrographs were discarded. Particles were picked in RELION-2.1

427   using reference classes obtained from a manually-picked, 50-micrograph dataset. Two

428   rounds of 2D classification were performed to discard poorly averaging particles. 3D

429   classification was performed using a 50 Å, low pass filtered initial model, based on results

430   from an *ab initio* reconstruction derived from cryoSPARC[63]. For the Widom 601 NCP,

431   particles contributing to 3D classes with discernible secondary-structure features were

432   pooled and refined using a spherical mask, and postprocessed in RELION-2.1[64] resulting in

433   a 3.8 Å (C1 symmetry applied) or 3.5 Å resolution (C2 symmetry applied). For the D02-NCP-

434   Streptavidin, a relatively smaller percentage of particles contributed to subnanometre-

435   resolution 3D averages. This is likely because of evident flexibility of the both the exit

436   nucleosomal DNA and the streptavidin group. To help drive streptavidin alignment and avoid

437   artificial NCP symmetrisation, a loose mask was used in a subsequent round of 3D

438   classification, encompassing both NCP and streptavidin. The resulting asymmetric

439   reconstruction yielded a reconstruction with 4.6 Å (no mask) or 4.2 Å resolution (loose

440   mask applied during refinement).

441

Atomic model docking and refinement

443   For the NCP-intasome STC complex NCP (3UTB[65] from PDBredo) and PFV strand transfer

444   complex (3OS0[16]) crystal structures were docked in the EM map using Chimera[66] and

445   clashing DNA segments were removed from the model. In order to refine the voxel/pixel size

446   of the map a series of maps were calculated with voxel/pixel size from 0.9 to 1.15 in steps of

447   0.01 and the initial model was refined against each map using phenix.real_space_refine[67]

448   with no additional geometry restraints. The geometry of resulting models was compared, and

449   voxel/pixel size fine-tuned between 1.11 -1.12 in steps of 0.001. The model refined against

450   the map with voxel/pixel size of 1.111 maintained the best geometry and was used for

451   further model building and refinement. The model was adjusted, and sequence of protein

452   and DNA components matched to the biological sample manually in Coot[68] and refined

453   using phenix.real_space_refine (Nightly build version 1.10pre-2091)[69] and Namdinator[70,71].

454   Additional restraints describing protein secondary structure, DNA base pairing and stacking

455   were used in Phenix. Protein geometry was assessed with Molprobity[67] and DNA geometry

456   was assessed with 3DNA[72]. For the D02 structure, NCP structure 5MLU was used as the

457   starting model to be independent from the NCP-intasome STC structure. The sequence was

458   adjusted and model manually tweaked in Coot and refined using phenix.real_space_refine

459   (Nightly build version phenix-dev-3374). Fine tuning of the voxel/pixel size was deemed

460   unnecessary as the model refined without issue. Both models have reasonable

461   stereochemistry and are in good agreement with the EM density maps.

462

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, 7K MWCO) to remove beta-mercaptoethanol. Octamers at 70 µM (140 µM of cysteine) were incubated with 5 mM TCEP for 10 minutes at room temperature. To achieve partial labelling, sulpho-Cy5 maleimide was added at 105 µM for 1 hour at room temperature. The reaction was quenched by adding 5mM beta-mercaptoethanol and desalted to remove unreacted dye (ThermoFisher, 7K MWCO). The extent of labelling was quantified by measuring the 595nm/280nm absorbance ratio, as well as by 2D intact mass ESI mass spectrometry. 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_r = 5,000, Laysan Bio, Inc.) containing 10% biotin-PEG-SVA (M_r = 5,000, Laysan Bio, Inc.) in 100 mM sodium bicarbonate, and used to construct a microfluidic channel as described previously[73]. 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 µM $ZnCl_2$ 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) microscope[74]. 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. For each experimental condition,

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

500

501   **Intasome strand-transfer and pull-down assays**

502   Intasome integration assays were performed as described[15], briefly 5 μg of NCPs were
503   incubated with 1.5 μg of intasome in intasome reaction buffer with and without 5mM $MgCl_2$
504   at 37°C for 15 minutes. The reaction was quenched by the addition of 25 mM EDTA and
505   0.2% SDS, and DNA precipitated after proteinase K digestion. DNA was then separated on
506   4-12% TBE polyacrylamide gels. Pull-down assays were performed as previously
507   described[15].

508

509

Authors contributions

510

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

517

518

532

533

Data availability

534

535   Model coordinates for the NCP-D02-streptavidin and Intasome-NCP structures are
536   deposited in the Protein Data Bank under accession code 6RNY and 6R0C respectively.
537   Cryo-EM maps for NCP-D02-streptavidin, NCP-601 and Intasome-NCP are available at the
538   EMDB under codes EMD-4692, EMD-4693 and EMDB-4960 respectively.

539
540

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).

594   24.   Mackler, R.M. et al. Nucleosome DNA unwrapping does not affect prototype foamy
595   virus integration efficiency or site selection. *PLoS One* **14**, e0212764 (2019).
596   25.   Wilson, M.D. & Costa, A. Cryo-electron microscopy of chromatin biology. *Acta*
597   *Crystallogr D Struct Biol* **73**, 541-548 (2017).
598   26.   Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W. & Richmond, T.J. Solvent
599   mediated interactions in the structure of the nucleosome core particle at 1.9 a
600   resolution. *J Mol Biol* **319**, 1097-113 (2002).
601   27.   Bilokapic, S., Strauss, M. & Halic, M. Histone octamer rearranges to adapt to DNA
602   unwrapping. *Nat Struct Mol Biol* **25**, 101-108 (2018).
603   28.   Lowary, P.T. & Widom, J. New DNA sequence rules for high affinity binding to
604   histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* **276**, 19-
605   42 (1998).
606   29.   Wilson, M.D. et al. The structural basis of modified nucleosome recognition by
607   53BP1. *Nature* **536**, 100-3 (2016).
608   30.   Chua, E.Y. et al. 3.9 A structure of the nucleosome core particle determined by
609   phase-plate cryo-EM. *Nucleic Acids Res* **44**, 8013-9 (2016).
610   31.   Iwasaki, W. et al. Comprehensive structural analysis of mutant nucleosomes
611   containing lysine to glutamine (KQ) substitutions in the H3 and H4 histone-fold
612   domains. *Biochemistry* **50**, 7822-32 (2011).
613   32.   McGinty, R.K. & Tan, S. Nucleosome structure and function. *Chem Rev* **115**, 2255-
614   73 (2015).
615   33.   McGinty, R.K., Henrici, R.C. & Tan, S. Crystal structure of the PRC1 ubiquitylation
616   module bound to the nucleosome. *Nature* **514**, 591-6 (2014).
617   34.   Makde, R.D., England, J.R., Yennawar, H.P. & Tan, S. Structure of RCC1 chromatin
618   factor bound to the nucleosome core particle. *Nature* **467**, 562-6 (2010).
619   35.   Arnaudo, N. et al. The N-terminal acetylation of Sir3 stabilizes its binding to the
620   nucleosome core particle. *Nat Struct Mol Biol* **20**, 1119-21 (2013).
621   36.   McGinty, R.K. & Tan, S. Recognition of the nucleosome by chromatin factors and
622   enzymes. *Curr Opin Struct Biol* **37**, 54-61 (2016).
623   37.   Zhou, K., Gaullier, G. & Luger, K. Nucleosome structure and dynamics are coming of
624   age. *Nat Struct Mol Biol* (2018).
625   38.   Anderson, C.J. et al. Structural Basis for Recognition of Ubiquitylated Nucleosome by
626   Dot1L Methyltransferase. *Cell Rep* **26**, 1681-1690 e5 (2019).
627   39.   Worden, E.J., Hoffmann, N.A., Hicks, C.W. & Wolberger, C. Mechanism of Cross-talk
628   between H2B Ubiquitination and H3 Methylation by Dot1L. *Cell* **176**, 1490-1501 e12
629   (2019).
630   40.   Valencia-Sanchez, M.I. et al. Structural Basis of Dot1L Stimulation by Histone H2B
631   Lysine 120 Ubiquitination. *Mol Cell* (2019).
632   41.   Jang, S. et al. Structural basis of recognition and destabilization of the histone H2B
633   ubiquitinated nucleosome by the DOT1L histone H3 Lys79 methyltransferase. *Genes*
634   *Dev* (2019).
635   42.   Bilokapic, S., Strauss, M. & Halic, M. Structural rearrangements of the histone
636   octamer translocate DNA. *Nat Commun* **9**, 1330 (2018).
637   43.   Kujirai, T. et al. Structural basis of the nucleosome transition during RNA polymerase
638   II passage. *Science* **362**, 595-598 (2018).
639   44.   Ehara, H. et al. Structural insight into nucleosome transcription by RNA polymerase II
640   with elongation factors. *Science* **363**, 744-747 (2019).
641   45.   Liu, X., Li, M., Xia, X., Li, X. & Chen, Z. Mechanism of chromatin remodelling
642   revealed by the Snf2-nucleosome structure. *Nature* **544**, 440-445 (2017).
643   46.   Farnung, L., Vos, S.M., Wigge, C. & Cramer, P. Nucleosome-Chd1 structure and
644   implications for chromatin remodelling. *Nature* **550**, 539-542 (2017).
645   47.   Eustermann, S. et al. Structural basis for ATP-dependent chromatin remodelling by
646   the INO80 complex. *Nature* **556**, 386-390 (2018).
647   48.   Ayala, R. et al. Structure and regulation of the human INO80-nucleosome complex.
648   *Nature* **556**, 391-395 (2018).

649 49.   Sundaramoorthy, R. et al. Structure of the chromatin remodelling enzyme Chd1
650   bound to a ubiquitinylated nucleosome. *Elife* **7**(2018).
651 50.   Sundaramoorthy, R. et al. Structural reorganization of the chromatin remodeling
652   enzyme Chd1 upon engagement with nucleosomes. *Elife* **6**(2017).
653 51.   Yan, L., Wu, H., Li, X., Gao, N. & Chen, Z. Structures of the ISWI-nucleosome
654   complex reveal a conserved mechanism of chromatin remodeling. *Nat Struct Mol Biol*
655   (2019).
656 52.   Li, M. et al. Mechanism of DNA translocation underlying chromatin remodelling by
657   Snf2. *Nature* **567**, 409-413 (2019).
658 53.   Lesbats, P. et al. Functional coupling between HIV-1 integrase and the SWI/SNF
659   chromatin remodeling complex for efficient in vitro integration into stable
660   nucleosomes. *PLoS Pathog* **7**, e1001280 (2011).
661 54.   Dyer, P.N. et al. Reconstitution of nucleosome core particles from recombinant
662   histones and DNA. *Methods Enzymol* **375**, 23-44 (2004).
663 55.   Li, X. et al. Electron counting and beam-induced motion correction enable near-
664   atomic-resolution single-particle cryo-EM. *Nat Methods* **10**, 584-90 (2013).
665 56.   Sorzano, C.O. et al. XMIPP: a new generation of an open-source image processing
666   package for electron microscopy. *J Struct Biol* **148**, 194-204 (2004).
667 57.   Scheres, S.H. RELION: implementation of a Bayesian approach to cryo-EM structure
668   determination. *J Struct Biol* **180**, 519-30 (2012).
669 58.   Mindell, J.A. & Grigorieff, N. Accurate determination of local defocus and specimen
670   tilt in electron microscopy. *J Struct Biol* **142**, 334-47 (2003).
671 59.   Bai, X.C., Fernandez, I.S., McMullan, G. & Scheres, S.H. Ribosome structures to
672   near-atomic resolution from thirty thousand cryo-EM particles. *Elife* **2**, e00461 (2013).
673 60.   Rosenthal, P.B. & Henderson, R. Optimal determination of particle orientation,
674   absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol*
675   *Biol* **333**, 721-45 (2003).
676 61.   Scheres, S.H. & Chen, S. Prevention of overfitting in cryo-EM structure
677   determination. *Nat Methods* **9**, 853-4 (2012).
678 62.   Zhang, K. Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12
679   (2016).
680 63.   Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for
681   rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296
682   (2017).
683 64.   Kimanius, D., Forsberg, B.O., Scheres, S.H. & Lindahl, E. Accelerated cryo-EM
684   structure determination with parallelisation using GPUs in RELION-2. *Elife* **5**(2016).
685 65.   Chua, E.Y., Vasudevan, D., Davey, G.E., Wu, B. & Davey, C.A. The mechanics
686   behind DNA sequence-dependent properties of the nucleosome. *Nucleic Acids Res*
687   **40**, 6338-52 (2012).
688 66.   Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research
689   and analysis. *J Comput Chem* **25**, 1605-12 (2004).
690 67.   Adams, P.D. et al. PHENIX: a comprehensive Python-based system for
691   macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21
692   (2010).
693 68.   Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of
694   Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
695 69.   Afonine, P.V. et al. Real-space refinement in PHENIX for cryo-EM and
696   crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544 (2018).
697 70.   Trabuco, L.G., Villa, E., Schreiner, E., Harrison, C.B. & Schulten, K. Molecular
698   dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and
699   X-ray crystallography. *Methods* **49**, 174-80 (2009).
700 71.   Kidmose, R.T. et al. Namdinator - Automatic Molecular Dynamics flexible fitting of
701   structural models into cryo-EM and crystallography experimental maps. *bioRxiv*
702   *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