Histone H1 binding to nucleosome arrays depends on linker DNA length and trajectory

Marco Dombrowski1, Maik Engeholm1, Christian Dienemann1, Svetlana Dodonova1,2
and Patrick Cramer1✉

Throughout the genome, nucleosomes often form regular arrays that differ in nucleosome repeat length (NRL), occupancy of linker histone H1 and transcriptional activity. Here, we report cryo-EM structures of human H1-containing tetranucleosome arrays with four physiologically relevant NRLs. The structures show a zig-zag arrangement of nucleosomes, with nucleosomes 1 and 3 forming a stack. H1 binding to stacked nucleosomes depends on the NRL, whereas H1 always binds to the non-stacked nucleosomes 2 and 4. Short NRLs lead to altered trajectories of linker DNA, and these altered trajectories sterically impair H1 binding to the stacked nucleosomes in our structures. As the NRL increases, linker DNA trajectories relax, enabling H1 contacts and binding. Our results provide an explanation for why arrays with short NRLs are depleted of H1 and suited for transcription, whereas arrays with long NRLs show full H1 occupancy and can form transcriptionally silent heterochromatin regions.
conformation in all our structures\textsuperscript{21}. Then, we used the overall EM maps to build the linker DNA connecting individual nucleosomes. This resulted in high-quality structures of the four arrays (Tables 1 and 2 and Supplementary Figs. 2–9). We could refine all four nucleosomes in the 4×177 array (Fig. 1c, Supplementary Figs. 2 and 3 and Supplementary Video 1) and could resolve the first three nucleosomes of the 4×187 (Supplementary Figs. 4 and 5 and Supplementary Video 2), 4×197 (Supplementary Figs. 6 and 7 and Supplementary Video 3) and 4×207 arrays (Supplementary Figs. 8 and 9 and Supplementary Video 4).

**Overall structure of tetranucleosome arrays.** All four structures show a zig-zag arrangement of nucleosomes (Fig. 2a), similar to what was observed in the 4×167 array crystal structure without H1 (ref. \textsuperscript{21}) and in designed nucleosome fibers\textsuperscript{22,26}. The overall architecture of all tetranucleosome arrays reported here is similar. In all structures, nucleosomes 1 and 3 form a canonical stack\textsuperscript{21}, whereas nucleosome 2 is located in a DNA loop between the two stacking nucleosomes and is rotated relative to the nucleosome stack (Fig. 2a). The distance between nucleosome 2 and the nucleosome stack increases with increasing NRL, which leads to increased mobility of nucleosome 2 (Supplementary Figs. 2, 4, 6 and 8). Nucleosome 4 is not stacked with nucleosome 2 and is increasingly mobile as the NRL increases. We were nevertheless able to refine the structure of nucleosome 4 as part of a tetranucleosome in the 4×177 array and also in isolation within the 4×187 array. The linker DNA connecting nucleosomes 3 and 4 was always visible and always showed the same trajectory as in the 4×177 structure.

**Fig. 1 | Reconstitution of tetranucleosome arrays for structural studies.** \textbf{a}, DNA templates contain four Widom-601 (ref. \textsuperscript{56}) nucleosome positioning sequences and variable linker DNA: 4×177 with 30-bp linker, 4×187 with 40-bp linker, 4×197 with 50-bp linker, and 4×207 with 60-bp linker. \textbf{b}, EMSA confirms that tetranucleosome arrays were reconstituted with saturating amounts of linker histone H1.4. Stoichiometry of H1 to nucleosome is denoted by H1:nuc. \textbf{c}, Structure of the 4×177 tetranucleosome array shows a zig-zag arrangement of nucleosomes, with nucleosomes 1 and 3 forming a stack and nucleosomes 2 and 4 extending from the stack. DNA is shown in gray and white, core histones in wheat, and H1 in purple.
Table 1 | Cryo-EM data collection, refinement and validation statistics for the 4×177 and 4×187 arrays

|      | nuc 1 | nuc 2 | nuc 3 | nuc 4 | stack | trinuc | tetranuc | nuc 1 | nuc 2 | nuc 3 | nuc 4 | stack | trinuc |
|------|-------|-------|-------|-------|-------|--------|----------|-------|-------|-------|-------|-------|--------|
|      | EMD-13359 | EMD-13360 | EMD-13361 | EMD-13362 | EMD-13358 | EMD-13357 | EMD-13356 | EMD-13369 | EMD-13368 | EMD-13367 | EMD-13366 | EMD-13365 | EMD-13363 |
| Data collection and processing |       |       |       |       |       |        |          |       |       |       |       |       |        |
| Magnification | ×81,000 | ×81,000 | ×81,000 | ×81,000 | ×81,000 | ×81,000 | ×81,000 | ×81,000 | ×81,000 | ×81,000 | ×81,000 | ×81,000 | ×81,000 |
| Voltage (kV) | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 |
| Electron exposure (e⁻/Å²) | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 |
| Defocus range (µm) | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 | 0.5–2.0 |
| Pixel size (Å/pix) | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 |
| Symmetry imposed | C₁ | C₁ | C₁ | C₁ | C₁ | C₁ | C₁ | C₁ | C₁ | C₁ | C₁ | C₁ | C₁ |
| Initial particle images | 1,341,160 | 1,341,160 | 1,341,160 | 1,341,160 | 1,341,160 | 1,341,160 | 1,341,160 | 1,259,654 | 1,259,654 | 1,259,654 | 1,259,654 | 1,259,654 | 1,259,654 |
| Final particle images | 174,476 | 128,860 | 174,476 | 20,621 | 174,476 | 174,476 | 20,621 | 110,706 | 61,926 | 110,706 | 51,385 | 110,706 | 27,515 |
| Map resolution (Å) | 4.6 | 5.1 | 4.5 | 79 | 6.0 | 7.2 | 9.5 | 4.0 | 3.8 | 4.0 | 3.8 | 4.0 | 3.8 |
| FSC threshold | | | | | | | | | | | | | |
| Map resolution range (Å) | 4.2–8.0 | 4.5–7.3 | 4.1–8.7 | 7.2–25 | 4.5–10.8 | 5.6–11 | 6.8–25 | 3.9–11.3 | 3.4–8.1 | 3.8–9.2 | 3.6–9.3 | 4.0–13.0 | 8.0–14 |
| Refinement | | | | | | | | | | | | | |
| Initial models used (PDB code) | 7K5Y | 7K5Y | 7K5Y | 7K5Y | 7K5Y | 7K5Y | 7K5Y | 7K5Y | 7K5Y | 7K5Y | 7K5Y | 7K5Y | 7K5Y |
| Model resolution (Å) | 4.6 | 4.6 | 5.1 | 5.1 | 4.5 | 4.5 | 79 | 79 | 6.0 | 6.0 | 7.2 | 7.2 | 9.5 | 9.5 |
| Model resolution range (Å) | 4.2–8.0 | 4.5–7.3 | 4.1–8.7 | 7.2–25 | 4.5–10.8 | 5.6–11 | 6.8–25 | 3.9–11.3 | 3.4–8.1 | 3.8–9.2 | 3.6–9.3 | 4.0–13.0 | 8.0–14 |
| Map sharpening B factor (Å²) | −150 | −520 | −150 | −300 | −310 | −300 | −500 | 0 | −50 | 0 | −50 | 0 | 0 |
| Model composition | | | | | | | | | | | | | |
| Non-hydrogen atoms | 13,304 | 13,304 | 13,880 | 13,880 | 14,058 | 14,058 | 26,403 | 26,403 | 40,119 | 40,119 | 54,040 | 54,040 | 13,470 | 13,470 |
| Protein residues | 768 | 768 | 843 | 843 | 843 | 843 | 1,536 | 1,536 | 2,379 | 2,379 | 3,222 | 3,222 | 843 | 843 |
| DNA | 352 | 352 | 354 | 354 | 342 | 342 | 364 | 364 | 694 | 694 | 1,040 | 1,040 | 843 | 843 |
| B factors (Å²) | | | | | | | | | | | | | |
| Protein | 224 | 224 | 140 | 140 | 252 | 252 | 252 | 252 | 203 | 203 | 248 | 248 | 180 | 180 |
| DNA | 294 | 294 | 191 | 191 | 315 | 315 | 419 | 419 | 372 | 372 | 220 | 220 | 190 | 190 |
| R.m.s. deviations | | | | | | | | | | | | | |
| Bond lengths (Å) | 0.006 | 0.005 | 0.006 | 0.006 | 0.006 | 0.006 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 |
| Bond angles (°) | 0.912 | 0.932 | 0.895 | 0.910 | 0.957 | 1.087 | 1.280 | 1.054 | 0.889 | 0.887 | 0.944 | 0.974 | 1.335 | 1.335 |
| Validation | | | | | | | | | | | | | |
| MolProbity score | 1.3 | 1.3 | 1.14 | 1.14 | 1.46 | 1.46 | 1.45 | 1.45 | 1.42 | 1.42 | 1.64 | 1.64 | 1.62 | 1.62 |
| Clashscore | 5.53 | 5.53 | 6.47 | 6.47 | 7.45 | 7.45 | 8.54 | 8.54 | 8.22 | 8.22 | 7.70 | 7.70 | 6.84 | 6.84 |
| Poor rotamers (%) | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

Nuc 1, 2 and 3 refer to nucleosome 1, 2 and 3; stack refers to stacked nucleosomes 1 and 3; trinuc refers to the trinucleosome consisting of nucleosomes 1, 2 and 3; tetranuc refers to the tetranucleosome.
Nucleosome stacking in solution. Previous work has revealed two main types of stacking interactions in nucleosome arrays.25,26. Type I interactions are closely packed stacks with contacts between H2A–H2B dimers, and have been observed in the crystal structure of the 4×197 array without H1 (ref. 23) and within the tetranucleosome units of the 12×177 and 12×187 cryo-EM structures (Fig. 2b). Type II interactions are more open, with slightly offset nucleosomes and the H4 N-terminal tail in close proximity to the acidic patch of the adjacent nucleosome, and have been observed in the 6×187 crystal structure with H1.

Table 2 | Cryo-EM data collection, refinement and validation statistics for the 4×197 and 4×207 arrays

|          | 4×197 |          | 4×207 |
|----------|-------|----------|-------|
| nuc 1    |       | nuc 2    |       |
| EMD-13372| PDB 7PFD | EMD-13373| PDB 7PFE |
| nuc 3    |       | stack    |       |
| EMD-13374| PDB 7PFF | EMD-13371| PDB 7PDF |
| trinuc   |       |          |       |
| EMD-13370| PDB 7PFA | EMD-13372| PDB 7PFW |
| nuc 1    |       | nuc 2    |       |
| EMD-13381| PDB 7PFE | EMD-13382| PDB 7PFV |
| nuc 3    |       | stack    |       |
| EMD-13383| PDB 7PF | EMD-13380| PDB 7PFU |
| trinuc   |       |          |       |
| EMD-13379| PDB 7PF | EMD-13371| PDB 7PF |

Data collection and processing

- **Magnification**: ×81,000
- **Voltage (kV)**: 300
- **Electron exposure (e−/Å²)**: 60
- **Defocus range (µm)**: 0.5–3.0
- **Pixel size (Å/pix)**: 1.05
- **Symmetry imposed**: C
- **Initial particle images**: 1,075
- **Final particle images**: 113,924
- **Map resolution (Å)**: 4.4
- **FSC threshold**: 0.143
- **Map resolution range (Å)**: 4.0–7.5
- **Refinement**
  - **Initial models used (PDB code)**: 7K5Y
  - **Model resolution (Å)**: 4.4
  - **Model resolution range (Å)**: 4.0–7.5
  - **Map sharpening B factor (Å²)**: -50
  - **Model composition**
    - Non-hydrogen atoms: 13,675
    - Protein residues: 843
    - DNA: 334
    - B factors (Å²)
      - Protein: 408
      - DNA: 421
  - **R.m.s. deviations**
    - Bond lengths (Å): 0.006
    - Bond angles (°): 0.915

Validation

- **MolProbity score**: 1.44
- **Clashscore**: 7.78
- **Poor rotamers (%)**: 0.0
- **Ramachandran plot**
  - Favored (%): 98.0
  - Allowed (%): 2.0
  - Disallowed (%): 0.0

NATURE STRUCTURAL & MOLECULAR BIOLOGY | VOL 29 | MAY 2022 | 493-501 | www.nature.com/nsmb
Fig. 2 | Structure of trinucleosome cores of tetranucleosome arrays. a. The trinucleosome cores of the 4×177, 4×187, 4×197 and 4×207 structures. Nucleosome 2 is rotated relative to the stack in all structures and is located at a greater distance from the stack as the length of linker DNA increases. Color code used throughout. b. Nucleosome stacking in nucleosome arrays. Nucleosome stacking in tetranucleosome arrays is similar to the stacking observed in the crystal structure of the 4×167 array without H1 (ref. 24) and the cryo-EM reconstruction of the 12×177 and 12×187 arrays with H1 (ref. 25). Left, nucleosome stack from the 4×187 array represents stacks from all structures reported in this study. Middle: nucleosome stack from the 4×167 crystal structure (PDB 1ZBB (ref. 23)) represents the type I interaction observed in the 4×167 crystal structure and within tetranucleosomal units of the 12×177 and 12×187 cryo-EM structures. Right, nucleosome stack from the 6×187 crystal structure (PDB 6HKT (ref. 26)) represents the type II interaction observed between tetranucleosome units of the 12×177 and 12×187 cryo-EM structures. Top, dyad axes drawn in green run almost parallel to the stacking observed in the cryo-EM reconstructions determined for 4×177, 4×187, 4×197 and 4×207, whereas dyad axes in the stack observed in both type I and type II interactions are slightly tilted toward each other. Bottom, the interface between stacking nucleosomes in the 4×177, 4×187, 4×197 and 4×207 structures reported here and in type I interactions consists of apposed H2a–H2B dimers (H2a in yellow, H2B in red), while in type II interactions the nucleosome stack is slightly offset and places the N-terminal part of H4 (green) near the H2A–H2B dimer.

Fig. 3 | NRL determines H1 binding to arrays. a. H1 binds to nucleosomes of the array near the nucleosome dyad. The N-terminal part of the α2-helix (Nα2) and the L3 loop contact the DNA around the dyad, whereas the α3-helix and the L1 loop interact with linker DNAs. H1 is rainbow-colored from the N (blue) to C (red) terminus, DNA is shown in white, and the histone octamer is shown in wheat. b. Focused-refined cryo-EM densities for nucleosomes 1, 2, 3 and 4, colored by NRL (4×177 blue, 4×187 green, 4×197 yellow, 4×207 red). H1 density is in purple. Nucleosomes are all viewed the same way. Entry and exit DNA are marked by a blue and a red dot, respectively. Focused-refined maps of nucleosome 4 could not be obtained for the 4×197 and 4×207 arrays owing to higher mobility. c. H1 N-terminal regions extend from the nucleosome stack in opposite directions. Residues regulating H1 mobility (K34 (ref. 34) and S35 (ref. 35)) and heterochromatin formation (K26 and S27) protrude from the nucleosome stack on both sides and are accessible for protein-protein interactions. The first ordered residue of H1 is S35; disordered residues are shown as a dashed line. DNA is shown in gray, histone octamer in wheat and H1 in purple.

In our structures, we observe a compact stacking of nucleosomes 1 and 3 that is similar to type I interactions with a contact formed between H2A–H2B dimers (Fig. 2b). The observed stacking does not allow for interactions between the H4 N-terminal tail of one
nucleosome with the acidic patch of a stacked nucleosome and thus leaves the H4 tail free to engage in other interactions\(^3\). Whereas the inter-nucleosome interactions appear to be very similar, we note a slight relative tilting of the stacking nucleosomes that positions their dyad axes almost parallel, in contrast to type I interactions in which the dyads are slightly tilted toward each other (Fig. 2b).

This difference might be due to the absence of H1 in the case of the 4x167 crystal structure\(^7\) and the different binding mode of H1 to the nucleosome in the case of the 12-mer array with H1 (ref. 20).

H1 orientation and DNA interactions. Our structures show that H1 is always bound near the nucleosome dyad (Fig. 3 and Supplementary Figs. 3, 5, 7 and 9). In all ten focused-refined maps, H1 shows three DNA contacts, similar to what has been described\(^12\). The H1 loop L3 and the N-terminal part of helix α2 contact nucleosomal DNA near the dyad, helix α3 binds one linker DNA and loop L1 contacts the other linker DNA (Fig. 3a and Supplementary Figs. 3, 5, 7 and 9). This mode of H1 binding is referred to as on-dyad\(^17\), although H1 is located slightly off the dyad and is lopsided\(^4\). H1 that is bound to nucleosome 1 always contacts entering linker DNA via its helix α3 (Fig. 3b, Supplementary Figs. 5, 7 and 9 and Supplementary Videos 2–4), whereas H1 on nucleosome 3 uses α3 to contact exiting linker DNA (Fig. 3b, Supplementary Fig. 9 and Supplementary Video 4).

In nucleosomes 2 and 4, the entering linker DNA is in contact with α3 (Fig. 3b and Supplementary Figs. 3 and 5; also see Supplementary Video 1).

Thus H1 can be oriented to either contact entering or exiting linker DNA, depending on local DNA geometry. The orientation of H1 influences the direction in which the unstructured N-terminal region of H1 exposes residues to post-translational modifications, such as K34 acetylation, S35 phosphorylation, K26 methylation and S27 phosphorylation\(^6\) (Fig. 3c). This places N-terminal H1 residues that have been shown to be important for either H1 mobility\(^34,35\) or heterochromatin formation\(^7\) at the surface of the nucleosome stack, where they are accessible to modifiers and binding partners even in the presence of a nucleosome stack.

H1 binding relates to nucleosome repeat length. The major difference between our four structures relates to the binding of the H1 histone to the different nucleosomes of the arrays (Fig. 3b). The H1 histone is present on nucleosome 2 in all four structures, and is also observed on nucleosome 4 in all cases where this nucleosome is structurally resolved. In contrast, the presence of H1 on the stacked nucleosomes 1 and 3 differs between the four arrays. H1 is absent from the stacked nucleosomes of the 4×177 array, but is present on nucleosome 1 in the 4×187 and 4×197 arrays, and is present on both stacked nucleosomes in the 4×207 array. Thus, histone H1 is bound to non-stacked nucleosomes in all structures, whereas H1 binding to stacked nucleosomes is enabled only as the NRL increases.

To confirm that our observations are not a result of low salt concentrations, we solved the trinucleosome core structure of the H1-bound 4×177 array at 60 mM NaCl and confirmed the presence of nucleosome stacks and the absence of H1 on stacking nucleosomes (Supplementary Fig. 11). We have also probed H1 binding to reconstituted 4×177, 4×187, 4×197 and 4×207 arrays biochemically at 150 mM NaCl and observed that the extent of H1 binding increased with increasing linker length, in line with our structural observations (Supplementary Fig. 12). In conclusion, an increase in NRL is related to stable binding of more H1 copies.

H1 binding depends on linker DNA trajectory. These observations suggested that linker DNA trajectory determines whether H1 can bind to nucleosomes within an array. We therefore analyzed the linker DNA trajectory at the entry and exit sites of the stacked nucleosomes in all structures. This analysis revealed a progressive change in the trajectory of linker DNA as the NRL increased (Fig. 4). To quantify this, we measured the angles α and β that define linker DNA geometry as described\(^6\) (Methods and Fig. 4b). Of particular importance here was angle β, formed between the nucleosome dyad and the linker DNA duplex axis. We also calculated the differences in angles, Δα and Δβ, which are the deviations between the angles α and β, respectively, observed in our structures and that in an isolated H1-bound nucleosome (PDB 7KSY (ref. 19)).

Our analysis showed that Δβ is a good predictor for histone H1 binding on stacked nucleosomes (Fig. 5). When Δβ was close to zero for both linker DNAs emerging from a nucleosome, H1 binding was observed (Fig. 5a). We found low Δβ values at nucleosome 2 and Δβ values of less than 6° at nucleosome 4, where H1 was always observed (Supplementary Table 1). However, when Δβ was higher, H1 was not bound, likely because a stabilizing contact between loop L1 and linker DNA could not be formed. Particularly high Δβ values are found for entry DNA at nucleosome 3, except for the 4×207 array, which is the only array where H1 is observed on nucleosome 3 (Fig. 5b).

Furthermore, exit DNA of nucleosome 1 shows the highest Δβ value for the 4×177 array, which is the only array in which H1 is lacking on this nucleosome (Fig. 5c). In summary, as the NRL increases, nucleosome 2 moves farther away from the stacked nucleosomes and the trajectories of linker DNA at nucleosomes 1 and 3 progressively approach canonical values (Δβ = ~0) (Fig. 5a). As a consequence, H1 can contact linker DNA, explaining H1 binding to stacked nucleosomes in arrays with longer NRLs (Fig. 6).

Discussion
We present cryo-EM structures of tetranucleosome arrays with different NRLs in the presence of the human linker histone H1 variant H1.4. The structures reveal a typical zig-zag arrangement of nucleosomes\(^21,26\), with a trinucleosome core consisting of two stacked nucleosomes 1 and 3 and a more flexible connecting nucleosome 2, suggesting that a trinucleosome may be a fundamental unit in chromatin\(^2\). The zig-zag arrangement is observed also in our 4×207 structure, in line with observations from in-cell mapping of DNA contacts\(^3\). Stacked nucleosomes have also been observed by structural studies of tetranucleosomes, trinucleosomes and free
mononucleosomes in solution\textsuperscript{32,39–42}. Stacking of nucleosomes 1 and 3 is apparently stabilized by H1 binding to nucleosome 2, because a published structure of a 3×177 trinucleosome array lacking H1 adopts a non-stacked, extended conformation\textsuperscript{41}. Our observation of a single nucleosome stack is consistent with small angle X-ray scattering (SAXS) analysis of tetranucleosomes\textsuperscript{42} and hexanucleosome arrays that showed limited compaction\textsuperscript{26}. Similar to previous structures of nucleosome arrays\textsuperscript{23,25,26}, the structures presented here use NRLs that correspond to those found in vivo\textsuperscript{7} and that differ by integer repeats of the approximate helical repeat of DNA (10n bp linkers with n being a natural number). However, alternative structures of trinucleosomes and tetranucleosomes certainly exist in vivo, and it will be important to study arrays with other linker lengths in the future\textsuperscript{43}.

Our major finding here is how the NRL of a nucleosome array relates to H1 binding to the array. It has long been known that there is a correlation between the NRL and the amount of associated H1 (refs. \textsuperscript{12,13}). Additionally, in vitro experiments showed that chromatin with closely spaced nucleosomes does not incorporate H1, whereas chromatin more widely spaced nucleosomes does\textsuperscript{44}, but the reasons for this remained elusive. We now report structures that show that short NRLs impair H1 binding (Supplementary Fig. 12) to stacked nucleosomes and suggest this is due to altered linker DNA trajectories. Altered linker DNA trajectories, as observed in our 4×177 array, sterically preclude H1-linker DNA contacts that are required for stable H1 binding\textsuperscript{17–19}. A similar observation was made in the structure of a nucleosome containing the H3 variant CENP-A, where an altered linker DNA trajectory has been observed\textsuperscript{41,45,46}. We show that, with increasing NRL, the linker DNA emerging from the stacked nucleosomes is more

![Fig. 5](image1.png)

**Fig. 5** | **Linker DNA trajectory determines H1 binding.** For each nucleosome, $\Delta \alpha$ and $\Delta \beta$ describe the difference in $\alpha$ and $\beta$, respectively, between isolated H1-bound mononucleosomal linker DNA (PDB 7K5Y (ref. \textsuperscript{19})) and the linker DNA of the nucleosomes in the tetranucleosome array (Supplementary Fig. 1). a, A plot of a nucleosome’s average $\Delta \alpha$ against its average $\Delta \beta$ reveals that nucleosomes not bound by H1 (ocher) separate well from the population of nucleosomes bound by H1 (purple). For nucleosome 3, they move closer to this population with increasing NRL. b, $\Delta \beta$ for nucleosome 3 entry DNA reveals a decrease with increasing NRL. c, $\Delta \beta$ for nucleosome 1 exit DNA reveals a decrease with increasing NRL. For the depicted nucleosomes, an overlay of the 4×177 nucleosome (blue) and the isolated H1-bound nucleosome (gray) is shown and $\Delta \beta$ for the different NRL arrays is listed, with bound H1 indicated by purple asterisks.

![Fig. 6](image2.png)

**Fig. 6** | **Overview of H1 binding to tetranucleosome arrays.** Note that H1 binding to stacked nucleosomes depends on linker DNA trajectory that in turn depends on the NRL. For details, compare text.
relaxed and permits stable H1 binding. Therefore, whereas H1 may transiently bind all nucleosomes of the four arrays (Fig. 1b), binding to nucleosomes might be destabilized in short NRL arrays and easily disrupted during cryo-EM sample preparation. We observe canonical on-dyad H1 binding as described6–19, in contrast to the off-dyad position of H1 found in tetranucleosome units of 12-mer arrays20 that is possibly a result of chemical crosslinking6.

Our results have important implications for understanding the relationship between the NRL of a genomic region and its transcriptional activity. In particular, the short NRls that are characteristic of active promoter regions and transcriptionally active gene bodies21,22 may preclude H1 from binding to stacked nucleosomes. This could explain the observed depletion of H1 from active promoters23,24,25 that likely facilitates assembly of the RNA polymerase II (Pol II) transcription machinery and passage of Pol II through chromatin6. The NRL of nucleosome arrays can be defined by chromatin remodeling enzymes26,27, and thus remodelers may indirectly deplete H1 by setting short NRls, thereby complementing other mechanisms of H1 depletion28,29 and rendering chromatin permissive to transcription.

Finally, long NRls are found in heterochromatin regions3,7,8,13, which seems counterintuitive because long NRls should expose more DNA to the transcription machinery but heterochromatin is transcriptionally silent. Our findings settle this apparent contradiction. We find that longer NRls are required to enable H1 binding to all nucleosomes of an array, thereby stabilizing nucleosomes and inhibiting chromatin remodeler activity10,30. Binding of H1 in turn widens the nucleosomal footprint against which remodelers move neighboring nucleosomes30–32 and thus would increase the NRL. Other H1-dependent mechanisms contribute to heterochromatin formation and transcriptional silencing33–35. For example, recruitment of DNA methyltransferases can downregulate transcription3, and heterochromatin protein 1 (HP1) binds to methylated H1 residue K26 (ref. 36) and may bridge H1-bound nucleosome stacks to facilitate heterochromatin formation and explain transcription repression.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41594-022-00768-w.

Received: 26 July 2021; Accepted: 1 April 2022; Published online: 17 May 2022

References
1. Baldi, S., Korber, P. & Becker, P. B. Beads on a string: nucleosome array arrangements and folding of the chromatin fiber. Nat. Struct. Mol. Biol. 27, 199–118 (2020).
2. Lai, W. K. M. & Pugh, B. F. Understanding nucleosome dynamics and their links to gene expression and DNA replication. Nat. Rev. Mol. Cell Biol. 18, 548–562 (2017).
3. Oruba, A., Saccone, S. & van Essen, D. Role of cell-type specific nucleosome positioning in inducible activation of mammalian promoters. Nat. Commun. 11, 1075 (2020).
4. Luger, K., Mäder, A. W., Richardson, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389, 251–260 (1997).
5. Baldi, S., Krebs, S., Blum, H. & Becker, P. B. Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. Nat. Struct. Mol. Biol. 25, 894–901 (2018).
6. Lai, B. et al. Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. Nature 562, 281–285 (2018).
7. Valouev, A. et al. Determinants of nucleosome organization in primary human cells. Nature 474, 516–520 (2011).
8. Chereji, R. V., Ramachandran, S., Bryson, T. D. & Henikoff, S. Nature-wide mapping of single nucleosomes and linkers in vivo. Genome Biol. 19, 19 (2018).
9. Frolov, D. V., Zhou, B. R., Skoultchi, A. I. & Bai, Y. Emerging roles of linker histones in regulating chromatin structure and function. Nat. Rev. Mol. Cell Biol. 19, 192–206 (2018).
10. Gunjan, A., Alexander, B. T., Sittman, D. B. & Brown, D. T. Effects of H1 histone variant overexpression on chromatin structure. J. Biol. Chem. 274, 37950–37956 (1999).
11. Fan, Y. et al. Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. Cell 123, 1199–1212 (2005).
12. Blank, T. A. & Becker, P. B. Electrostatic mechanism of nucleosome spacing. J. Mol. Biol. 252, 305–313 (1995).
13. Woodcock, C. L., Skoultchi, A. I. & Fan, Y. Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. Chromosome Res. 14, 17–25 (2006).
14. Shimada, M. et al. Gene-specific H1 eviction through a transcriptional activator→p300→NAP1→H1 pathway. Mol. Cell 74, 268–283.e5 (2019).
15. Dorigo, J., Chereji, R. V., Eriksson, P. R. & Clark, D. J. The ISW1 and CHD1 ATP-dependent chromatin remodelers compete to set nucleosome spacing in vivo. Nucleic Acids Res. 44, 4625–4635 (2016).
16. Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L. & Sweet, R. M. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. Nature 299, 223 (1993).
17. Zhou, B. R. et al. Structural mechanisms of nucleosome recognition by linker histones. Mol. Cell 59, 628–638 (2015).
18. Rednar, J. et al. Structure and dynamics of a 197 bp nucleosome in complex with linker histone H1. Mol. Cell 66, 384–397.e8 (2017).
19. Zhou, B. R. et al. Distinct structures and dynamics of chromatosomes with different human linker histone isoforms. Mol. Cell 81, 166–182.e6 (2021).
20. Zhou, B. R. & Bai, Y. Chromatin structures condensed by linker histones. Essays Biochem. 63, 75–87 (2019).
21. Chen, P., Li, W. & Li, G. Structures and functions of chromatin fibers. Annu. Rev. Biophys. 50, 95–116 (2021).
22. Dorigo, J. et al. Nucleosome arrays reveal the two-start organization of the chromatin fiber. Science 306, 1571–1573 (2004).
23. Schalch, T., Duda, S., Sargent, D. F. & Richmond, T. J. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature 436, 138–141 (2005).
24. Ekdunayo, B., Richmond, T. J. & Schalch, T. Capturing structural heterogeneity in chromatin fibers. J. Mol. Biol. 429, 3031–3042 (2017).
25. Song, F. et al. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science 344, 376–380 (2014).
26. García-Saez, I. et al. Structure of an H1-bound 6-nucleosome array reveals an untwisted two-start chromatin fiber conformation. Mol. Cell 72, 902–915 (2018).
27. Ou, H. D. et al. ChromEMT: visualizing 3D chromatin structure and compaction in interphase and mitotic cells. Science 357, eaag0025 (2017).
28. Cai, S., Böck, D., Pilhofer, M. & Gan, L. The in situ structures of mono-, di-, and trinucleosomes in human heterochromatin. Mol. Biol. Cell 29, 2450–2457 (2018).
29. Ricci, M. A., Manzo, C., Garcia-Parajo, M. F., Lakadamyali, M. & Cosma, M. P. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. Cell 160, 1145–1158 (2015).
30. Ohno, M. et al. Sub-nucleosomal genome structure reveals distinct nucleosome folding motifs. Cell 176, 520–534.e25 (2019).
31. Tsunaka, Y., Kajimura, N., Tate, S. & Morikawa, K. Alteration of the nucleosomal DNA path in the crystal structure of a human nucleosome core particle. Nucleic Acids Res. 33, 3424–3434 (2005).
32. Bilokapic, S., Strauss, M. & Halic, M. Cryo-EM of nucleosome core particle interactions in trans. Sci. Rep. 8, 7046 (2018).
33. Kan, P. Y., Caterino, T. L. & Hayes, J. T. The H4 tail domain participates in intra- and inter-nucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays. Mol. Cell 59, 538–546 (2009).
34. Kamieniarz, K. et al. A dual role of linker histone H1.4 Lys 34 acetylation in transcriptional activation. Genes Dev. 26, 797–802 (2012).
35. Chu, C. S. et al. Protein kinase A-mediated serine 35 phosphorylation associates histone H1.4 from mitotic chromosomes. J. Biol. Chem. 286, 35843–35851 (2011).
36. Daujat, S., Zeissler, U., Waldmann, T., Happel, N. & Schneider, R. HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. J. Biol. Chem. 280, 38090–38095 (2005).
37. Ishihara, S. et al. Local states of chromatin compaction at transcription start sites control transcription levels. Nucleic Acids Res. 49, 8007–8023 (2021).
38. Risca, V. I., Denny, S. K., Straight, A. F. & Greenleaf, W. J. Variable chromatin structure revealed by in situ spatially correlated DNA cleavage mapping. Nature 541, 237–241 (2017).
39. Mauney, A. W., Muthurajan, U. M., Luger, K. & Pollack, L. Solution structure(s) of trinucleosomes from contrast variation SAXS. *Nucleic Acids Res.* **49**, 5028–5037 (2021).
40. Ding, X., Lin, X. & Zhang, B. Stability and folding pathways of tetra-nucleosome from six-dimensional free energy surface. *Nat. Commun.* **12**, 1091 (2021).
41. Takizawa, Y. et al. Cryo-EM structures of centromeric tri-nucleosomes containing a central CENP-A nucleosome. *Structure* **28**, 44–53.e4 (2020).
42. Zhou, B. R. et al. Revisit of reconstituted 30-nm nucleosome arrays reveals an ensemble of dynamic structures. *J. Mol. Biol.* **430**, 3093–3110 (2018).
43. Zhurkin, V. B. & Norouzi, D. Topological polymorphism of nucleosome fibers and folding of chromatin. *Biophys. J.* **120**, 577–585 (2021).
44. Lusser, A., Urwin, D. L. & Kadonaga, J. T. Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* **12**, 160–166 (2005).
45. Zhou, B. R. et al. Atomic resolution cryo-EM structure of a native-like CENP-A nucleosome aided by an antibody fragment. *Nat. Commun.* **10**, 2301 (2019).
46. Roulland, Y. et al. The flexible ends of CENP-A nucleosome are required for mitotic fidelity. *Mol. Cell* **63**, 674–685 (2016).
47. Izzo, A. et al. The genomic landscape of the somatic linker histone subtypes H1.1 to H1.5 in human cells. *Cell Rep.* **3**, 2142–2154 (2013).
48. Krishnakumar, R. et al. Reciprocal binding of PARP-1 and histone H1 at promoters specifies transcriptional outcomes. *Science* **319**, 819–821 (2008).
49. Krietenstein, N. et al. Genomic nucleosome organization reconstituted with pure proteins. *Cell* **167**, 709–721.e12 (2016).
50. Maier, V. K., Chioda, M., Rhodes, D. & Becker, P. B. ACF catalyses chromatosome movements in chromatin fibres. *EMBO J.* **27**, 817–826 (2008).
51. Oberbeckmann, E. et al. Ruler elements in chromatin remodelers set nucleosome array spacing and phasing. *Nat. Commun.* **12**, 5232 (2021).
52. Yamada, K. et al. Structure and mechanism of the chromatin remodelling factor ISW1a. *Nature* **472**, 448–453 (2011).
53. Eggers, N. & Becker, P. B. Cell-free genomics reveal intrinsic, cooperative and competitive determinants of chromatin interactions. *Nucleic Acids Res.* **49**, 7620–7647 (2021).
54. Yang, S. M., Kim, B. J., Norwood Toro, L. & Skoultschi, A. I. H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation. *Proc. Natl Acad. Sci. USA* **110**, 1708–1713 (2013).
55. Sun, J. et al. Histone H1-mediated epigenetic regulation controls germline stem cell self-renewal by modulating H4K16 acetylation. *Nat. Commun.* **6**, 8856 (2015).
56. 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).

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

Articles

molar ratio 1.2:1.2:1:1, dialyzed 3 times against gel filtration buffer (20 mM HEPES pH 7.5, 10 mM dithiothreitol (DTT)), core histones were mixed at −80 °C. The sample was adjusted to 8 M urea by weighing in solid urea, added 200 mM HEPES pH 7.5 and run over a His-Trap 1 ml HP (GE Healthcare) column. Full-length Smt3-H1.4-GyrA were cleaved by Ulp1 for 1 hour at room temperature, Rosetta 2 (DE3) and incubated on ice for 30 min, and binding was probed by EMSA as described.

Nucleosome array reconstitution. Nucleosome arrays containing H1.4 were reconstituted by salt-gradient dialysis as described. Briefly, histone octamer and nucleosome reconstitution buffer A (20 mM HEPES pH 7.0, 2 M NaCl, 1 mM EDTA, 1 mM DTT) transferred into Slide-A-Lyzer Mini Dialysis Units 3.500 MWC COOH (Thermo Scientific) filter cups and gradually dialyzed over 16 hours from nucleosome reconstitution buffer A to nucleosome reconstitution buffer B (20 mM HEPES pH 7.0, 600 mM NaCl), concentrated using Amicon Ultra-4 10kDa MWCO centrifugal filters (Merck Millipore) and directly used for nucleosome reconstitution or were flash-frozen in liquid nitrogen and stored at −80°C.

Nucleosome arrays were reconstituted with H1.4 and analyzed by Banl restriction enzyme digestion. For EMSAs of H1-containing arrays, 300 ng of sample was run on a 1.2% agarose 0.5xTBE buffer for 1.5 hours at 110 V at 4°C. To test differential binding of histone octamer to the Widom-601 nucleosome positioning sequence, nucleosome arrays were reconstituted without H1.4 and analyzed by Banl restriction enzyme digestion. For EMSAs of H1-containing arrays, 300 ng of sample was run on a 1.2% agarose 0.5xTBE buffer for 1.5 hours at 110 V at 4°C. To test differential binding of H1.4 to arrays of different NRLLs, nucleosome arrays were reconstituted in the absence of H1.4 and adjusted to 100 mM DNA and 150 mM NaCl. H1.4 was then added to different molar ratios of H1 to Widom-601 sequence and incubated on ice for 30 min, and binding was probed by EMSA as described above. For sample in buffer with salt, the sample was adjusted to 60 mM NaCl and incubated for 30 min on ice prior to cryo-EM grid preparation.

Cryo-EM sample preparation and data collection. Quantifoil Cu 300 B 1.2/1.3 holey carbon grids were glow-discharged using a PELOQ easiGlow (Ted Pella) for 100s at 15 mA and 0.4 bar. In a Vitrobot Mark IV (FEI) chamber set to 100% humidity and −20°C for 20 s, the liquid was blotted away using blot force 5 for 3 seconds, and the grid was vitrified by plunging into liquid ethane. Data were collected on a Titan Krios 300 kV transmission electron microscope (FEI) equipped with a Gatan Imaging Filter set to 20 e− and a K3 direct electron detector (Gatan). Movies containing 60 frames with a total fluence of 60 e−/Å2 were collected using SerialEM at a nominal magnification of ×8,100 and a pixel size of 1.05 Å/pixel with 40° stage tilt.

Data processing and analysis. Gain normalization, motion correction and CTF estimation of cryo-EM movies were performed using Warp, and particles were picked using an instance of Warp's neural network retrained on the 4×177 data set. Particles were extracted at 8.4 Å/pixel in RELION 3.1 (refs. and ) and sorted by 2–3 rounds of two-dimensional classification in cryoSPARC. Particles belonging to classes showing 2 or more nucleosome cores were reconstructed at 3.15 Å/pixel, and all subsequent processing was done in RELION 3.1.

For the 4×177+H1.4 data set (Supplementary Fig. 2), several rounds of 3D classification yielded particles that were refined to a 7.2-Å resolution map of a 4×177 trimucleosome. From this, 3D classification with a mask around the presumed location of the nucleosome 4 yielded particles that were refined to a 9.5-Å resolution map of the 4×177 tetranucleosome. The signal of the trimucleosome was subtracted from these particles, and the output was refined to the 7.9-Å resolution map of the fourth nucleosome. From the 4×177 trimucleosome map, masked refinements on the nucleosome stack or the connecting nucleosome were signal subtracted for the other nucleosomes and refined to yield the focused-refined maps of nucleosomes 1, 2, and 3.

Similarly, the 4×187 (Supplementary Fig. 4), 4×197 (Supplementary Fig. 6) and 4×207 (Supplementary Fig. 8) cryo-EM data were subjected to several rounds of 3D classification and 3D refinement to yield maps with a defined nucleosome stack and blurred density for the connecting nucleosome. From this map, several more rounds of 3D classification were used to select particles. These particles were refined to the 4×187, 4×197 and 4×207 trimucleosome at 11 Å, 9.7 Å and 9.8 Å resolution, respectively. Particles from the 3D refinement of the stack with less defined connecting nucleosomes were extracted, unbinned and further processed using signal subtraction, 3D classifications and masked refinements to yield maps for nucleosome data set. For the 4×177 data set, the selected particles were refined to the 4×177+H1.4 and 4×207+H1.4 data sets. The angular distribution of views for each map was plotted using Warp, local resolution and global FSC was determined using RELION, and the directional FSCs were calculated using the 3D FSC server.

Model building and refinement. The local-resolution-filtered maps were used for model building, except for the 4×177 trimucleosome, 4×177 nucleosome 1, 4×177 nucleosome 2 and 4×177 nucleosome 4, for which the post-processed maps were used. For each data set, the structure of the H1-bound mononucleosome (PDB 7K51 (ref. ) with protein and DNA sequences mutated to the ones used in this study, was rigid-body fitted into the density of nucleosomal unit in UCSF Chimera. Protein termini, entry DNA and exit DNA were manually adjusted in COOT, and the resulting structures were real-space refined in PHENIX. The refined nucleosomal units were then rigid-body fitted into corresponding densities of the nucleosome stack, trimucleosome and tetranucleosome, respectively, using UCSF Chimera. In case of the trimucleosome and tetranucleosome structures, the linker DNA was manually built in COOT. The models were real-space refined in PHENIX and were validated using Molprobity (Tables 1 and 2). Figures were generated using PyMOL (Schrodinger), UCSF Chimera and UCSF ChimeraX.

Analysis of linker DNA trajectories. The models for the nucleosome stacks were used to measure linker DNA trajectories for nucleosomes 1 and 3, and the models of the focused-refined maps of nucleosomes 2 and 4 were used to measure linker DNA deviation for nucleosomes 2 and 4. The corresponding maps were used to rigid-body fit the structure of the H1-bound 197 bp mononucleosome (PDB 7K51 (ref. )) into the density of the nucleosome disc. Linker vectors u, v and w were calculated using (4) the centroid of coordinates of the base pair 5 bp into the Widom-601 sequence. (ref. ). The plane of the nucleosome disc needs to be defined to determine the angle u by using (4) the centroid of coordinates of the base pair 5 bp into the Widom-601 sequence and u point on the plane and the centroid of points 2 and 3 to approximate the normal to the nucleosome disc, and v point on the plane and the centroid of points 2 and 3 to approximate the dyad axis. We used u and v to describe the plane perpendicular to the disc nucleosome. We determined the normal w to this plane by taking the normalized cross product of u and v, and we use u and v and w to describe the plane of the nucleosome disc. Linker vectors were defined by using (4) the centroid of coordinates of the base pair 5 bp into the Widom-601 sequence and (5) the centroid of the coordinates of the base pair 10 bp outside of the Widom-601 sequence. For measurement of the angle β, as shown in Fig. 6b, we projected linker DNA vectors onto the plane generated by u and v and calculated the angle between the projected vectors. For the angle α, linker DNA vectors were projected onto plane the plane generated by u and w and we calculated the angle between the projected vectors. Calculations were done in MATLAB R2017a.

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

Data availability. Electron microscopy densities have been deposited in the EM Data Bank with the accession codes EMD-13359 to EMD-13383. The coordinate files have been deposited in the Protein Data Bank with the accession codes 7PEW to 7PFX. See Tables 1 and 2. Source data are provided with this paper.

NATURE STRUCTURAL & MOLECULAR BIOLOGY | www.nature.com/nsmb
References

57. Dodonova, S. O., Zhu, F., Dienemann, C., Taipale, J. & Cramer, P. Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function. *Nature* **580**, 669–672 (2020).

58. Osunsade, A. et al. A robust method for the purification and characterization of recombinant human histone H1 variants. *Biochemistry* **58**, 171–176 (2019).

59. Lis, J. T. & Schleif, R. Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucleic Acids Res.* **2**, 383–389 (1975).

60. Luger, K., Reichsteiner, T. J. & Richmond, T. J. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol.* **304**, 3–19 (1999).

61. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).

62. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019).

63. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 319–530 (2012).

64. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

65. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

66. Tan, Y. Z. et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).

67. Petsersen, E. E. et al. UCSF Chimera — a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

68. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

69. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

70. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).

Acknowledgements

We thank E. Oberbeckmann, M. Ochmann and H. Wang for critical reading of the manuscript. S.D. was supported by an EMBO long-term fellowship (ALTF 949-2016). M.E. was supported by the Deutsche Forschungsgemeinschaft (EXC 2067/1-390729940). P.C. was supported by the Deutsche Forschungsgemeinschaft (EXC 2067/1-390729940) and the European Research Council Advanced Investigator Grant CHROMATRANS (grant agreement No. 882357).

Author contributions

M.D. designed and conducted all experiments and data analysis unless stated otherwise. M.E. prepared cryo-EM samples of the 4×177 array. C.D. carried out cryo-EM data acquisition for the 4×177 array, maintained the EM facility and advised on microscope setup. S.D. conducted image processing of the 4×177 array data. S.D. and P.C. initially outlined the project. P.C. supervised research. M.D., S.D. and P.C. wrote the manuscript, with input from M.E. and C.D.

Funding

Open access funding provided by Max Planck Society.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41594-022-00768-w.

Correspondence and requests for materials should be addressed to Svetlana Dodonova or Patrick Cramer.

Peer review information Nature Structural and Molecular Biology thanks the anonymous reviewers for their contribution to the peer review of this work. Beth Moorefield was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.
Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a  Confirmed

☑ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement

☑ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

☑ The statistical test(s) used AND whether they are one- or two-sided

**Only common tests should be described solely by name; describe more complex techniques in the Methods section.**

☑ A description of all covariates tested

☑ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

☑ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

☑ For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted

*Give P values as exact values whenever suitable.*

☑ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

☑ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

☑ Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

*Our web collection on statistics for biologists contains articles on many of the points above.*

Software and code

Policy information about availability of computer code

| Data collection | SerialEM 3.8 beta 8 |
|-----------------|---------------------|

| Data analysis   | RELION 3.1.0:commit:3b3228, UCSF Chimera 1.9, UCSF ChimeraX 0.91, Pymol 2.2.2, Warp v1.0.7, Phenix 1.19.2, Coot 0.9, cryoSPARC v2.14.2, MATLAB R2017a, 3DFSC, MolProbity 4.5 |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Cryo-EM reconstructions and atomic coordinates have been deposited via the wwPDB OneDep system. Corresponding PDB and EMDB IDs are found in Tables 1 and 2. Where publicly available data has been used, ID has been stated in the manuscript.
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size: Micrographs have been collected until ca 1 million particles have been autopicked by Warp for each data set.

Data exclusions: No data excluded from analysis.

Replication: Cryo-EM averages over large number of independent observations.

Randomization: n.a.

Blinding: n.a.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

- n/a
- Involved in the study
- □ Antibodies
- □ Eukaryotic cell lines
- □ Palaeontology and archaeology
- □ Animals and other organisms
- □ Human research participants
- □ Clinical data
- □ Dual use research of concern

### Methods

- n/a
- Involved in the study
- □ ChiP-seq
- □ Flow cytometry
- □ MRI-based neuroimaging