学位論文

Structural basis of CENP-SX chromatin and the stability of FANCM-CENP-SX complex

（CENP-SXクロマチンの構造基盤と
FANCM-CENP-SX複合体の安定性について）

2021年3月

東京理科大学大学院　基礎工学研究科　生物工学専攻

伊藤 翔

Contents

Abstract

Faithful chromosomal segregation and maintenance of genome stability are indispensable for cell functions and the perpetuation of life. CENP-SX that is a conserved histone-fold complex can bind to double stranded DNA and functions in both kinetochore assembly and DNA repair. Chromosomal segregation requires a protein complex known as kinetochore that mediates attachment between mitotic spindle microtubules and centromere-specific nucleosomes. CENP-SX is one of the components and form complex with another histone-fold complex CENP-TW. In DNA repair, on the other hand, CENP-SX interacts with FANCM, one of the FA (Fanconi Anemia) causative genes, to enhance the FANCM stability and activity. CENP-TWSX and FANCM-CENP-SX have distinct DNA-binding properties to play important roles in kinetochore assembly and DNA repair, respectively. However, how these complexes are organized and functions on DNA is not fully understood.

Here, I aimed to understand how CENP-SX binds to DNA. Using chicken recombinant proteins expressed in *Escherichia coli*, I investigated the CENP-SX DNA-binding activity and discovered the CENP-SX DNA-binding mode that has not been well characterized. Structural analysis showed the CENP-SX-DNA unit is an asymmetric structure that three CENP-SX tetramers oligomerize on DNA and covers about 22 bp linear DNA. The unit explains the biochemical characterization of CENP-SX regular DNA-binding ability and suggests a unique chromosomal structure formed by CENP-SX which I term 'CENP-SX chromatin'. Further biochemical and structural analysis shows FANCM, a partner of CENP-SX in DNA repair, can be incorporated in CENP-SX chromatin, which suggests FANCM-CENP-SX chromatin. Investigating the FANCM-CENP-SX complex stability indicates FANCM is released under the condition that antagonizes the interaction between FANCM and CENP-SX, which remains CENP-SX chromatin intact. In conclusion, I revealed the structural basis of CENP-SX DNA binding characteristics and provide a new DNA-binding mode by histone-fold protein, which suggests CENP-SX chromatin structure as a scaffold for other protein complexes.

Introduction

Eukaryotic DNA is wrapped around histones to form nucleosomes and chromatin. Replication, transcription, and repair machinery transverses nucleosomes with the help of remodeling complexes and maintains genome integrity (*1*). During mitosis, chromosomes are condensed, and they segregate to two daughter cells at the onset of anaphase through the attachment of the spindle microtubule. Spindle microtubules attach to the kinetochore built at the centromeric chromatin. Many protein complexes are involved in these processes (*2*).

CENP-SX is a histone-fold complex conserved from budding yeast to human and is known to function both in chromosome segregation and in DNA repair (*3–7*). In chromosome segregation, it forms a complex with another kinetochore histone-fold complex CENP-TW to form CENP-TWSX heterotetramer (*4*). CENP-TWSX hetero-tetramer complex can bind DNA and introduce positive supercoil into DNA while CENP-SX, CENP-TW, and canonical histones introduce negative supercoil. This suggests centromere-specific CENP-TWSX chromatin but its detail remains elusive. On the other hand, It is reported that CENP-SX has a role in limiting the crossover formation at meiosis which is distinct from kinetochore function (*8–10*). It also functions to repair inter-strand crosslinks formed by DNA damaging agent (*11*). In this pathway, CENP-SX forms a complex with FANCM (*6, 7*). FANCM is a component of the Fanconi anemia DNA-repair system (*12*). FANCM also functions independent of the FA repair pathway; stabilizing and remodeling replication fork, and limiting crossovers (*13–21*). FANCM and its homologs possess SF2 helicase domain and CENP-SX binding motif (*22–24*). CENP-SX stabilizes the FANCM structure and enhances its DNA binding activity. Previous studies have revealed that CENP-SX facilitates the action of the bound partner by strengthening DNA binding (*6*, *7*, *25*, *26*). Several structures of human FANCM-CENP-SX complexes are known, and they form heteropentamers (*25, 26*). FANCM-binding is thought to alter CENP-SX DNA-binding preference from dsDNA to branched DNA but how the FANCM-CENP-SX complex recognizes DNA substrates is still unclear.

CENP-SX has the ability to condense nucleosome and promote annealing of complementary DNA strands (*7*), which could assist the FANCM activity at the DNA damage site. Furthermore, CENP-SX possesses unique DNA binding property distinct from CENP-TW or FANCM. Electrophoresis Mobility Shift Assay (EMSA) shows that CENP-SX forms a ladder-like regular band shift pattern (*4, 6, 7*), suggesting that it forms a regular unit similar to nucleosomes. However, the CENP-SX-DNA complex is not as stable as the nucleosome and is easily processed by micrococcal nuclease (*4*). Regular DNA binding pattern requires both the histone-fold domain and CENP-S intrinsic disordered region (IDR) containing basic residues (*4, 7, 25*). However, how

CENP-SX forms regular chromatin structure remains elusive. In this study, I investigated the DNA recognition mechanisms of histone-fold complex CENP-SX by structural and biochemical analyses. As a consequence, I propose 'CENP-SX chromatin' that is distinct from that of the canonical histone-based nucleosome. In addition, although I've not obtained the FANCM-CENP-SX-DNA crystals, the FANCM-CENP-SX complex could adopt the same chromatin structure as CENP-SX by the superimposition of newly determined chicken FANCM-CENP-SX structure to CENP-SX chromatin. These results suggest that CENP-SX chromatin offers a platform for the complex formation and function with other interacting partners.

Results

CENP-S IDR property determines CENP-SX DNA binding profile

To understand the DNA binding property of CENP-SX, I made a careful sequence analysis of the CENP-S IDR (Figure 1A). Although the sequence identity is low, the electrostatic pattern was relatively conserved. A conserved basic region was present just after the histone-fold which was followed by a conserved acidic region. To test the significance of these regions, I examined the DNA binding ability by comparing the DNA binding of the full-length CENP-SX (CENP-SX), CENP-SX without the acidic region (CENP-S$^{\Delta Ac}$X), and CENP-SX without IDR (CENP-S$^{\Delta IDR}$X). Since the previous study showed CENP-SX strongly binds to more than 40 bp dsDNA (*4*), I started with 49 bp dsDNA (Figure 1B). As shown previously, CENP-S$^{\Delta IDR}$X does not form a regular ladder-like band shift pattern. On the other hand, CENP-S$^{\Delta Ac}$X forms a band similar to CENP-SX. This suggests that the basic region but not the acidic region within the CENP-S IDR is required to form a stable DNA binding complex. Interestingly, the shifted band of CENP-S$^{\Delta Ac}$X was more smeared compared to the CENP-SX.

To analyze the binding property of CENP-SX in more detail, I carefully compared the DNA binding by changing the CENP-SX tetramer-DNA ratio using 49 bp dsDNA (Figure 1B). In CENP-SX, the amount of shift band increased gradually from the ratio of 1:1 and 1:2. In this condition, a free DNA band was still present. However, at a ratio of 1:3, the shifted band in CENP-SX increased dramatically and the free DNA completely disappeared at a ratio of 1:4. On the other hand, CENP-S$^{\Delta Ac}$X formed a gradual band shift and it lacked a dramatic increase in DNA binding at a ratio of 1:3 or 1:4. When the bound complex and protein/DNA ratio were quantified, CENP-SX showed a steep inflection point whereas CENP-S$^{\Delta Ac}$X lacking this property (Figure 1C). These results suggest that CENP-SX tetramers show a cooperative DNA binding property which is controlled by basic and acidic regions within the CENP-S IDR.

Having identified the DNA property of CENP-SX binding to 49bp dsDNA, I next used other dsDNA lengths to see whether there is a preference in the cooperativity and binding pattern (Figure 1B). Using dsDNA length ranging from 19 bp to 97 bp, CENP-SX formed a regular band shift pattern and the number of bands increased as the length of DNA increased, as reported previously. I observed a single band within 37-55 bp dsDNA, two bands within 61-73 bp dsDNA, and three bands with 79-97 bp. Intriguingly, within 19-30 bp, the amount of free DNA apparently decreased at a ratio of 1:4-1:6, whereas a discrete ladder band could not be detected. Comparing the band intensity from 19 to 55 bp indicated that the intensity increased with the lengths. Since Ultra Power DNA staining dye binds to the exposed hydrophobic surface, it is possible that 19-30 bp DNA in

complex with CENP-SX were incompetent for visualization. Taken together, these results suggest CENP-SX could form a unit with about 30 bp DNA. In all cases, band shift by CENP-S$^{\Delta Ac}$X was more smeared and there was more free dsDNA.

CENP-SX chromatin structure determination and its characteristics

To elucidate the structural basis of the CENP-SX-DNA complex, I tried to crystallize CENP-SX in complex with DNA. Optimizing the crystallization condition (protein/DNA ratio, DNA length, precipitant reagent concentration, etc.), I was successful to crystallize CENP-SX with 30 or 31 bp DNA reproducibly and obtained two crystal forms ($P2_1$ and C2) from the similar crystallization condition. One crystal belonged to the space group $P2_1$ and diffracted to 3.6 Å in resolution. Another crystal belonged to the space group C2 and diffracted to 3.2 Å (Table 2). Both structures were solved by molecular replacement method using CENP-SX and dsDNA. $P2_1$ crystal was refined to Rwork/Rfree values of 24.9%/29.3% and C2 crystal was refined to Rwork/Rfree values of 23.1%/26.1%. After solving the structures, I found that both were very similar with a root mean square displacement (rmsd) of 1.24 Å. C2 asymmetric unit content was half of the $P2_1$ and related by crystallographic symmetry (Figure 2A). Electron density of CENP-S 6-115 aa, CENP-X 6-80 aa, 22bp dsDNA were observed in the $P2_1$ structure, and CENP-S 8-115 aa, CENP-X 6-80 aa, 18bp dsDNA were observed in the C2 structure.

Since the shape and electrostatic surface pattern of CENP-SX are similar to other histone-fold proteins, the bound DNA trajectory on the CENP-SX tetramer was predicted to be similar as well. However, what was striking was that single dsDNA was encircled by three CENP-SX tetramers forming a chromatin unit (Figure 2B). Each CENP-SX tetramer bound to about 16 bp DNA and neighboring tetramer bound to the next 16 bp with 3 bp step size along the DNA helix (Figure 2C right). In total, three CENP-SX tetramers covered linear 22 bp DNA. DNA was relatively straight with no obvious deformation or bends when compared to the dsDNA structure.

The electrostatic surface of three CENP-SX tetramers showed their histone-fold (HF) DNA-binding interfaces formed the basic cavity that fit the phosphate backbones of the DNA (Figure 2B, C), indicating that basic residues of this interface recognize the backbone and facilitate three CENP-SX tetramers' arrangement on DNA. This is consistent with the previous data that alanine exchange for the basic residue of this interface abolished ladder bands, and instead, formed a smear-like feature on EMSA (*4*).

In addition to HF DNA-binding interfaces, two CENP-S α4 helices from each CENP-SX tetramer sandwich the DNA with point symmetry (Figure 2C right, 2D). CENP-S C-terminal

positively charged region formed extended α-helix structure providing the other DNA binding surfaces (Figure 2D). The participation of this region in DNA binding consistent with the previous low-resolution CENP-SX-DNA crystal structure (*27*). In the absence of DNA, this region was unstructured and electron density was not observed. Thus, DNA binding induced the conformational change. On the other hand, the electron density of the C-terminal negatively charged region was still missing in the crystal structure possibly due to the structural flexibility. Although this region was unstructured, given the result from the EMSA (Figure 1B), C-terminal negatively charged region may be important for CENP-SX oligomers to bind dsDNA.

The extended helix region provides not only additional DNA-binding interface but also protein-protein interactions observed only when CENP-SX oligomerized on DNA. The three tetramers formed a specific interaction between the neighboring molecules. The extended CENP-S α4 helix of a tetramer formed electrostatic interaction with the adjacent CENP-X α1 helix (Figure 3A). In particular, CENP-S (S102, E106) and CENP-X (R15, R18) and additional van der Waals contacts stabilized tetramer-tetramer interaction. To see if the interaction affects CENP-SX oligomerization state on DNA, I introduced a single mutation that could defy the electrostatic interaction into these residues. CENP-SX^{R15A}, CENP-SX^{R18A}, CENP-S^{E106A}X could afford to form the ladder bands, whereas CENP-SX^{R15E}, CENP-SX^{R18E}, CENP-S^{E106R}X gave a smear-like band on EMSA (Figure 3B left). This smear-like band was partly restored by introducing further mutation that could recover the electrostatic interaction (CENP-S^{E106R}X^{R15E}, CENP-S^{E106R}X^{R18E}) (Figure 3B right). These results indicate that the specific oligomerization of the CENP-SX tetramers enables them to bind along the right-handed double helix.

Three tetramers are related by almost 90˚ with the first and the third tetramer on the opposite side of the DNA helix. When I tried to model in the fourth tetramer on either side, the preoccupied tetramer prevented the placement. Moreover, there is a disordered acidic region after the newly structured basic region, and this could also block the additional interaction. Thus, three CENP-SX tetramers and a dsDNA likely complete the CENP-SX-DNA complex which I term ‘CENP-SX chromatin’.

FANCM$^{660-804}$-CENP-SX-DNA complex formation

The structure of CENP-SX chromatin provides the basis for complementary DNA annealing and recombination. Moreover, there may be a unique function when CENP-SX and FANCM form a complex. To explore how CENP-SX chromatin is involved in DNA repair, I co-expressed and purified recombinant chicken FANCM-CENP-SX consisting of chicken FANCM that interacts

with CENP-SX (referred to as FANCM$^{660-804}$; amino acids 660-804), full-length of CENP-S and CENP-X (Figure 4A). The purified FANCM$^{660-804}$-CENP-SX complex was subjected to EMSA by mixing with 19-49 bp dsDNA to check FANCM$^{660-804}$-CENP-SX DNA-binding ability. It bound to all dsDNA length by adding an excess amount of protein (Figure 4B). After confirming its ability, I carefully examined whether FANCM-binding to CENP-SX affects binding pattern by changing the protein-DNA ratio using 37 bp (Figure 4C). FANCM$^{660-804}$-CENP-SX formed a regular protein-DNA complex similar to that of CENP-SX but the property was somewhat different. Consistent with the crystal structure and previous results, CENP-SX DNA-binding was associated with protein concentration to DNA mainly indicating a single ladder band with 37 bp. On the other hand, at a ratio of 1:1 and 1:2 (DNA:protein), FANCM$^{660-804}$-CENP-SX gave a ladder band at a lower position compared to that of CENP-SX. However, at a ratio of 1:3, another ladder band appeared at the almost the same position as that of CENP-SX-DNA. Furthermore, at a ratio of 1:3-1:6, FANCM-CENP-SX formed another band at the higher position than those. These results suggest that FANCM$^{660-804}$-CENP-SX can bind to DNA with lower stoichiometry than CENP-SX and change its oligomerization state with an increase in protein ratio. Altogether, these results may reflect distinct FANCM-CENP-SX-DNA complex formation which is different from CENP-SX alone.

FANCM$^{660-804}$-CENP-SX-DNA crystallization results in CENP-SX chromatin

To reveal the nature of the complex of FANCM-CENP-SX and DNA, I tried to crystallize FANCM$^{660-804}$-CENP-SX with dsDNA. FANCM$^{660-804}$-CENP-SX was successfully co-crystallized with 19-49 bp mixing the equal amount of crystallization reagent (Figure 5A). During the crystallization, there was a heavy film-like structure formed at the air-liquid interface of the crystal droplet. Optimizing crystallization and cryo-preservation condition, I obtained several crystals with diffraction quality sufficient for data processing. To determine the crystal structure, the molecular replacement was performed using chicken CENP-SX and linear dsDNA as a template. I was able to obtain phase and the calculated electron density map showed FANCM-missing CENP-SX chromatin structure that is almost similar to that of the CENP-SX dsDNA structure (Figure 5B). Indeed, when crystals in the crystallization drop were dissolved in the solution and analyzed by SDS-PAGE, only CENP-S and CENP-X bands were detected, whereas the film-like structure covering the droplet contained FANCM$^{660-804}$, CENP-S, and CENP-X (Figure 5C). I could not detect free FANCM$^{660-804}$ released from the droplet. It was surprising that FANCM$^{660-804}$-CENP-SX disassembled and CENP-SX formed co-crystals with dsDNA similar to the CENP-SX-DNA

structure. This suggests that a small fraction of CENP-SX might have been present in the purified complex or that components in the crystallization condition may have disrupted the complex.

The FANCM$^{660-804}$ release under the crystallization condition

To analyze the tripartite complex stability in more detail, I prepared chicken FANCM-CENP-SX complex without CENP-S IDR (FANCM$^{660-804}$-CENP-S$^{\Delta IDR}$X) and performed crystallization experiments. The purified complex was crystallized under various conditions and three different crystals were obtained (Figure 6A top). In the crystal droplet, as in the case of the complex crystallization with dsDNA, a heavy film-like structure formed at the air-liquid interface. Interestingly, structural determinations showed one of the crystals contained the FANCM$^{660-804}$-CENP-S$^{\Delta IDR}$X complex but the other two contained the CENP-S$^{\Delta IDR}$X complex without FANCM$^{660-804}$ (Figure 6A bottom, Table 3). Similar to the DNA complex crystallization, FANCM$^{660-804}$ existed in the film-like structure together with CENP-S$^{\Delta IDR}$X and the droplet did not contain FANCM$^{660-804}$ (Figure 6B).

To further assess the nature and stability of the FANCM$^{660-804}$-CENP-S$^{\Delta IDR}$X complex, I performed Native-PAGE (Figure 7Aa). Comparison of FANCM$^{660-804}$-CENP-S$^{\Delta IDR}$X (left) and CENP-S$^{\Delta IDR}$X (right) indicates that both complexes migrate as a discrete band and the former complex migrates faster than the latter. I confirmed the content by cutting out each band and analyzed them by SDS-PAGE (Figure 7Ab). The purified FANCM$^{660-804}$-CENP-S$^{\Delta IDR}$X seems to be nearly homogenous as judged by the gel filtration peak (Figure 7B) and the major fraction was indeed FANCM$^{660-804}$-CENP-S$^{\Delta IDR}$X (Figure 7Ab, *1 and *2), however, there was a small amount (~3%) of CENP-S$^{\Delta IDR}$X without FANCM$^{660-804}$ (Figure. 7Ab, *3). Interestingly, when I added 2-methyl-2,4-pentanediol (MPD) which was included in the protein crystallization regents, the band pattern did not change until 20% MPD but a further increase to 30% MPD resulted in the formation of a smeared band, which migrate to a position similar to CENP-S$^{\Delta IDR}$X. CENP-S$^{\Delta IDR}$X did not form such a band even at 30% MPD which suggests that MPD promotes dissociation of FANCM$^{660-804}$ from the complex. Other components in the crystallization buffer such as carboxylic acid mix did not cause such an effect (data not shown). As the film formed during the crystallization, which was likely caused by the exposure to the air, I speculated that oxidation and MPD promote distortion and dissociation of the complex.

Sequence analysis of the chicken FANCM$^{660-804}$ indicates that there are four cysteine residues within the current construct and two of them are conserved between chicken and human FANCM (Figure 7C). Cys 759 is buried within the complex and may not be sensitive to oxidation (Figure

7D). On the other hand, Cys670, Cys679, and Cys779 are exposed to solvent and in particular, Cys670 is in the disordered region and may be sensitive to oxidation. Cys679 and Cys779 are located within the alpha-helix and may be less sensitive. Thus, the truncation of the disordered region in the current chicken FANCM construct might increase the stability of the complex. As for the effect of MPD, it is a commonly used organic solvent in crystallography due to its amphiphilic nature and small flexible structure that binds to many parts of the protein (*28*). FANCM$^{660\text{-}804}$ interacts with CENP-SX through the extensive surface area of more than 5000 Å^2. The interaction is formed by hydrogen bonding, salt bridges, and hydrophobic interaction. Mapping of CENP-SX$^{\Delta IDR}$ hydrophobicity indicates that the interface is more hydrophobic than the other regions (Figure 7D). Therefore, the MPD might bind and promote the release of FANCM$^{660\text{-}804}$. Released FANCM$^{660\text{-}804}$ on its own is unstable and is prone to aggregation and binds to other FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X complexes through its hydrophobic surface and cysteine disulfide bridges.

FANCM-CENP-SX chromatin models

On the other hand, the FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X structure could be superimposed to the CENP-SX chromatin structure without steric hindrance (Figure 8). Combining the results of FANCM$^{660\text{-}804}$-CENP-SX EMSA, another model that two FANCM$^{660\text{-}804}$-CENP-SX complex bind to a single dsDNA could be attained because the moiety of electrostatic interactions enabling CENP-SX chromatin is retained in its model. As reported before, CENP-SX heterotetramer can exchange its dimer to CENP-TW dimer to form CENP-TWSX heterotetramer (*4*) and the moiety of electrostatic interactions between molecules may be insufficient for maintaining its oligomerization state. On the contrary, FANCM-binding stabilizes CENP-SX tetramerization and does not allow the release of free CENP-SX, and could bind DNA in the lower stoichiometry state. Collectively, these models can explain the FANCM$^{660\text{-}804}$-CENP-SX regular DNA-binding profile observed in EMSA, providing a model of FANCM-CENP-SX chromatin structure.

Discussion

Since its discovery, how CENP-SX recognizes DNA and facilitates kinetochore assembly and/or DNA repair has been a critical question. Based on the structural similarity, CENP-SX DNA binding mode was predicted to be similar to other histone-fold proteins including canonical histones. However, current biochemical and structural analyses revealed a novel DNA binding mode that is distinct from previous studies or other predictions. Three CENP-SX heterotetramers cooperatively oligomerized on 22 bp linear DNA to form an asymmetric functional unit. This unit was established by following three properties. First, the CENP-S C-terminal region has unique roles; CENP-S C-terminal positively charged region is essential for interacting with DNA and its extension of the negatively charged region gives cooperativity to promote and stabilize CENP-SX oligomerization on DNA. Second, the basic cavity formed by the arrangement of CENP-S loops belonging to HF DNA-binding interface fits the DNA backbone and allows CENP-SX to align along the DNA helix. Third, protein-protein electrostatic interactions between CENP-SX heterotetramers stabilize oligomerized state on DNA. This DNA binding mode is different from the previously reported low-resolution CENP-SX-DNA complex structure which bound DNA through the heterodimer and the basic region (*27*) (Figure 9). The oligomeric binding mode might have provided CENP-SX to bind relatively tightly on sequence-independent DNA.

CENP-SX participate in higher order chromatin structure formation but its exact nature has been unknown (*7*). A recent study showed the linker DNA of tri-nucleosome is 22-30 bp DNA (*29*). CENP-SX chromatin can be integrated into the model, to support linker dynamics. Further study using nucleosome should uncover the CENP-SX chromatin function incorporating into the canonical histone-based chromatin.

CENP-SX and CENP-TW form the CENP-TWSX complex that is thought to introduce centromere-specific chromatin to centromeric DNA for a kinetochore scaffold (*4*, *30*). The previous report demonstrates CENP-TW strongly binds to dsDNA but without a regular DNA-binding pattern. However, the addition of CENP-SX promotes CENP-TWSX complex formation and provides distinct regular DNA-binding similar to CENP-SX. Based on the CENP-SX chromatin structure, two distal CENP-SX dimers on either side could be replaced by CENP-TW dimer without structural disturbances. Previous biochemical analyses suggest relatively a plastic nature of the CENP-TWSX-DNA complex and I propose mosaic CENP-SX and CENP-TWSX chromatin (Figure 10). Further analysis of the CENP-TWSX-DNA complex in relation to CENP-SX will shed light on the nature of its DNA-binding mode.

It is reported that CENP-SX promotes complementary DNA strand annealing. CENP-SX chromatin structure in complex with FANCM might provide the molecular basis of its activity at the damaged DNA site occurring at the replication fork, D-loop, and so on. Here, I hypothesized that the new chromosomal architecture 'CENP-SX chromatin' is responsible for DNA-binding and processing activity of the FANCM-CENP-SX complex and tried to elucidate how CENP-SX chromatin is engaged in genome maintenance. DNA-binding assay and structure determination of FANCM-CENP-SX suggest the formation of FANCM-CENP-SX chromatin. Although I have not obtained the crystal structure of the FANCM-CENP-SX-DNA complex yet, I found that CENP-SX chromatin was stable under the absence of FANCM. Structural analysis of the FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X complex suggests it is sensitive to oxidation and the environment that defies the hydrophobic interaction resulting in FANCM$^{660\text{-}804}$ dissociation. Reflecting on the FANCM$^{660\text{-}804}$-CENP-SX-DNA crystallization condition, 1,4-Dioxane is also a commonly used organic compound and a high concentration of 1,4-Dioxane can alter protein structure (*31*). It also functions as an oxidant reagent and may promote protein aggregation by undesired disulfide bridge through the flexible cysteine residue. Thus, FANCM$^{660\text{-}804}$ was likely to be dissociated from CENP-SX under the crystallization condition, where CENP-SX chromatin maintained its structure. These results suggest FANCM inactivation doesn't perturb the CENP-SX chromatin structure.

In the future, investigating CENP-SX chromatin in the context of full-length FANCM will be the next step toward a holistic understanding of synergistic functions of the FANCM-CENP-SX complex and its chromatin assembly.

Conclusion

CENP-SX is involved in both kinetochore assembly and DNA repair and has been investigated by both research areas about its DNA-binding ability because of its sequential and structural resemblance to canonical histone proteins. In this study, I elucidated CENP-SX DNA-binding property by structural and biochemical analysis and proposed CENP-SX chromatin that is a distinct protein-DNA architecture from canonical nucleosomal chromatin. The structural basis of CENP-SX chromatin provides new insights into kinetochore formation and limiting crossover mechanism, suggesting that the novel chromosomal architecture 'CENP-SX chromatin' may be incorporated in canonical histone-based biological processes.

Methods

Protein secondary structure prediction

CENP-S secondary structure prediction was performed by Psipred server (*32–34*) and the disordered prediction plots were obtained (see Figure 1A).

Protein expression

The recombinant plasmid encoding 6×His-tev-CENP-S, and strep Ⅱ-tev-CENP-X was cloned in BLSS *Escherichia coli* cells (BL21Star (Thermofisher) with pRARE2LysS plasmid (Novagen)). For EMSA, CENP-S was truncated at residue E 106 (termed CENP-$S^{\Delta IDR}$) and A118 (termed CENP-$S^{\Delta Ac}$). Recombinant cells were grown in LB containing 50 μg/mL kanamycin at 37°C until OD600 reached 0.5. The protein expression was induced by 0.1 mM IPTG and the culture was incubated at 20°C for 12-15 hours.

The chicken FANCM gene was cloned from the chicken DT40 cDNA library. The recombinant plasmid encoding MBP-6×His-TEV protease recognition site (abbreviated as tev)-FANCM (660-804 aa, 6×His-tev-CENP-S, and strepII-tev-CENP-X was used to transform *Escherichia coli* cells (BL21Star (Thermofisher) with pRARE2LysS plasmid (Novagen)). For protein crystallization, CENP-S was truncated at residue E106 (termed CENP-$S^{\Delta IDR}$). Recombinant cells were grown in LB or Terrific Broth containing 100 μg/mL ampicillin at 37°C until OD_{600} reached 0.7-1.0. The protein expression was induced by 0.2 mM IPTG and the culture was incubated at 20°C for 12-15 hours.

The cells were harvested by centrifugation using JLA8.1000A rotor (BECKMAN) at 4,000 rpm for 15 minutes at 4°C and the bacterial pellet was stored at -80°C.

Protein purification

CENP-SX, CENP-$S^{\Delta Ac}$X, CENP-$S^{\Delta IDR}$X bacterial pellet were thawed and resuspended in buffer containing 10 mM Tris-HCl pH 8.0, 500mM NaCl. Resuspended cells were lysed by sonication using sonicator (ASTRASON, MISONIX, power of 8, Pulsar on for 0.2 seconds, Pulsar off for 0.8 seconds, Elapsed Time for 3-5 minutes), and cell debris was pelleted by centrifugation using R18A rotor (BECKMAN) at 11,500 rpm for 15 minutes at 4°C. The supernatant was applied to a 5-10 mL Histrap FF crude column (GE Healthcare) pre-equilibrated with 10 mM Tris-HCl pH 8.0, 500mM NaCl, 25 mM Imidazole, and bound proteins were eluted by 500 mM imidazole. The eluate was loaded onto a Superdex 200 pg column (GE Healthcare) equilibrated in 10 mM Tris-HCl pH 8.0, 500mM NaCl. Fractions containing target proteins were subjected to TEV cleavage reaction with the addition of homemade TEV protease and 1 mM dithiothreitol (DTT) at room temperature overnight. The solution was applied to Histrap FF crude column pre-equilibrated with 10 mM Tris-HCl pH 8.0, 500mM NaCl, 25 mM Imidazole. Flow-through fractions

were concentrated to 2 mL or 5 mL by ultrafiltration with Amicon Ultra MWCO 30,000 (Millipore) at 5,000 × g at 4°C and loaded onto a Superdex 200 pg column equilibrated in 10 mM Tris-HCl pH 8.0, 500mM NaCl. The peak fractions containing CENP-S and CENP-X were concentrated by ultrafiltration with Amicon Ultra MWCO 30,000 and stored at -25°C.

FANCM$^{660\text{-}804}$-CENP-SX and FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X also purified according to the same method as CENNP-SX. As for ultrafiltration, Amicon Ultra MWCO 50,000 (Millipore) was used.

Quality of the protein purification was analyzed by SDS-PAGE using 15% gel. The protein concentration was measured by UV absorption at 280 nm and calculated using extinction coefficient ε = 5960 / MW; 49358, ε = 5960 / MW; 44404, ε = 5960 / MW; 41484, ε = 32,680 / MW; 59683, and ε = 32,680 / MW; 70633 for CENP-SX, CENP-S$^{\Delta Ac}$X, CENP-S$^{\Delta IDR}$X, FANCM$^{660\text{-}804}$-CENP-SX, and FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X, respectively.

Preparation of DNA substrates

Complementary strands of oligonucleotides (at 500 μM each) (for sequence see Table 1) were mixed 1:1 and heated to 95°C for 10 minutes and then slowly cooled to 4°C. Annealed DNA was just used for the crystallization trial without further purification. For electron mobility shift assay (EMSA), annealed DNA was mixed with Ultra Power staining reagent, loaded onto TBE polyacrylamide gel, and electrophoresed. The gel containing DNA of interest was cut out then eluted by TE buffer.

Crystallization and cryopreservation

Crystallization of CENP-SX-dsDNA was set up in sitting-drop vapor diffusion format in MRC2 plates with drops comprised of 1 μL precipitant solution and 1 μL of protein-DNA complex and incubated at 20°C. Before the crystallization experiment, CENP-SX-DNA complex formation was performed in a solution containing 10 mM Tris-HCl pH 8.0, 100 mM NaCl with 150-200 μM protein, and 50 μM dsDNA for 30-60 minutes (see Table 1 for the DNA substrate used). Crystals appeared in a solution containing 50 mM MES-NaOH pH 6.5 or Bis-tris pH 6.0, 40% MPD, and 100 mM NaCl within a week and cryoprotected by using crystallization solution supplemented with 20% ethylene glycol and 20% DMSO.

Crystallization of FANCM$^{660\text{-}804}$-CENP-SX-dsDNA was conducted as follows. To form a protein-DNA complex, a mixture of 50 μM protein and 55 μM DNA (19-49 bp from widom 601, see Table 1) in solution (20 mM Tris-HCl pH 8, 100 mM NaCl, 5 mM 2-Mercaptoethanol) was incubated at 20°C for 60 minutes. Sitting drop trials were performed by mixing 1 μL of the protein-DNA and 1 μL precipitant containing 20 mM Tris-HCl pH 8, 100 mM NaCl, 31% 1,4-Dioxane, 10 mM $MgCl_2$, 5 mM 2-Mercaptoethanol. Crystallization plates were kept at 20°C and crystals appeared after 10-20 days. For cryopreservation,

harvest buffer (25 mM Tris-HCl pH 8, 110 mM NaCl, 12 mM $MgCl_2$, 5 mM 2-Mercaptoethanol, 30% MPD) was added to the crystallization drop and equal volume of harvest buffer was removed. This was repeated 2-3 times to fully exchange the buffer. Crystals were cryo-protected using a solution containing 25 mM Tris-HCl pH 8, 110 mM NaCl, 12 mM $MgCl_2$, 5 mM 2-Mercaptoethanol, 30% MPD, 10% DMSO or Ethylene Glycol.

The initial crystallization for FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X was carried out using a crystallization robot Mosquito (TTP Labtech). FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X was crystallized by mixing 100 nl of 10 mg/mL (176 µM) protein solution and 100 nl of Morpheus MD-HT screen buffer G12 (0.1 M Tris-Bicine pH 8.5, 0.1 M Carboxylic acids mix, 12.5% 2-methyl-2,4-pentanediol (MPD), 12.5% PEG1000, 12.5% PEG3350) (Molecular Dimensions). Needle and tetrahedral crystals appeared after 7 days. Crystals were soaked in the crystallization buffer for diffraction measurement. Manual crystallization was carried out by sitting drop crystallization at 20°C using 1 µL of concentrated protein solution and 1 µL of homemade crystallization reagent (0.1 M MOPS-HEPES pH 7.0, 0.1 M Carboxylic acids mix, 15% MPD, 20% PEG3350, 300 mM NaCl). Crystals appeared after 2-10 days. The crystals were transferred and soaked in 0.1 M MOPS-HEPES pH 7.0, 15% MPD, 25% PEG3350, and 350 mM NaCl supplemented with 0.6 M 1,6-Hexanediol for 2 hours.

Data collection and analysis

Diffraction data for CENP-SX-DNA were collected at BL1A of Photon Factory. Data were processed with XDS (*35*). Three CENP-SX heterodimers with a single-stranded DNA and three CENP-SX heterotetramers with a double-stranded DNA were used as search copies for C2 and $P2_1$ crystal, respectively. For the DNA component, an ideal B-form dsDNA was generated in Discovery Studio (Biovia) from the 25 bp DNA containing dyad sequence of nucleosome structure (PDB; 3LZ0).

Diffraction data for FANCM$^{660\text{-}804}$-CENP-SX-dsDNA were collected at BL44XU of Spring-8 and for FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X at BL1A and BL17A of Photon Factory and BL44XU of Spring-8. Data were processed with the HKL2000 package (HKL Research).

All molecular replacement was carried out using PHASER program of the PHENIX suite (*36*). All refinement was also carried out using PHENIX refinement mode and model building using COOT (*37*). Data collection and final refinement statistics are summarized in Tables 2 and 3. The structures were visualized with Discovery studio (Biovia).

Electrophoresis mobility shift assay (EMSA)

DNA-binding reactions were carried out for 30-60 minutes at room temperature in 10 µL solution containing 20 mM Tris-HCl pH 8, 100 mM NaCl, 5 mM 2-mercaptethanol with final concentration of 0.25, 0.5, 0.75,

1.0, 1,25, 1,5 μM CENP-SX to 0.25 μM DNA and 0.4, 0.8, 1.2, 1.6, 2.0, 2.4 μM FANCM$^{660\text{-}804}$-CENP-SX to 0.4 μM DNA. As for CENP-SX mutants, mixing 0.4, 1.2, 2.4 μM protein to 0.4 μM DNA. After the addition of 1 μL of 10× gel loading buffer (40% Glycerol, 0.2 M Tris-HCl pH 6.8, 1/300 Ultra Power (Gellex)), the reaction mixtures were applied to 6-9% native polyacrylamide gel in 0.5×TBE buffer and electrophoresis was performed at 4°C for 75-90 minutes at 120V. The gel was visualized under 500 nm LED light. Band intensities were quantified with ImageJ.

Native PAGE analysis

Protein solution containing 5 μL of 2 mg/ml FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X and 5 μL of A0 buffer (10 mM Tris-HCl pH 8), 10-60% MPD, or 10- carboxylic acids mix was incubated at 20°C for 60 min and then added by 2 μL of Native-PAGE loading buffer (0.25 M Tris-HCl pH 6.5, 20% sucrose, 0.02% BPB). 10 μL of the mix was loaded onto 6% polyacrylamide TBE (Tris, Boric acid, EDTA pH 8.0) gel. The gels were run at 150 V for 75-90 min in 0.5x TBE buffer at 4°C and subsequently stained with Coomassie brilliant blue R-250. bands migrated on TBE gel were clipped out, mashed, suspended with 1x SDS-PAGE loading buffer (0.05 M Tris-HCl pH 6.5, 4% sucrose, 0.005% BPB, 5% 2-ME, 2% SDS) and incubated at 94°C for 5 min. The supernatant was loaded onto an SDS-PAGE gel.

Figures & Tables

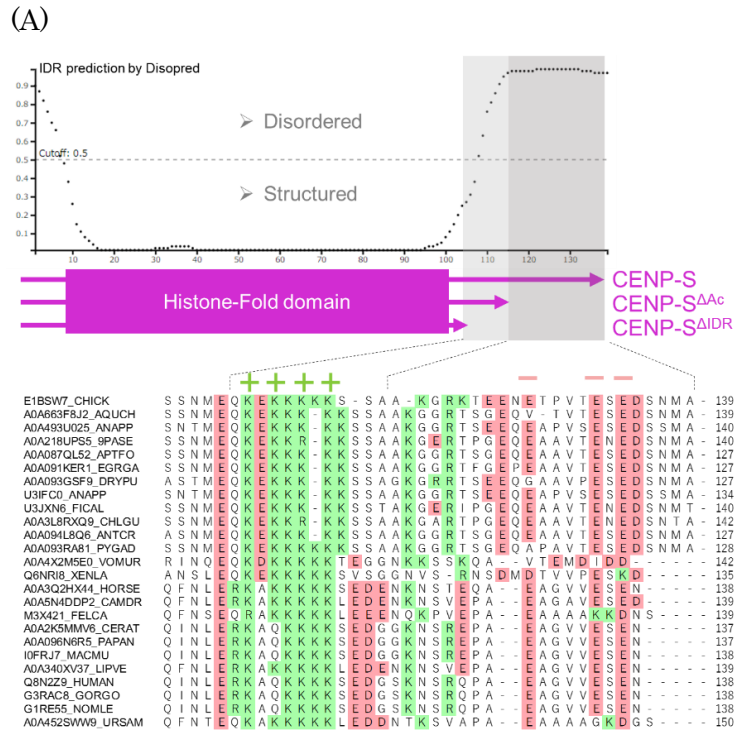
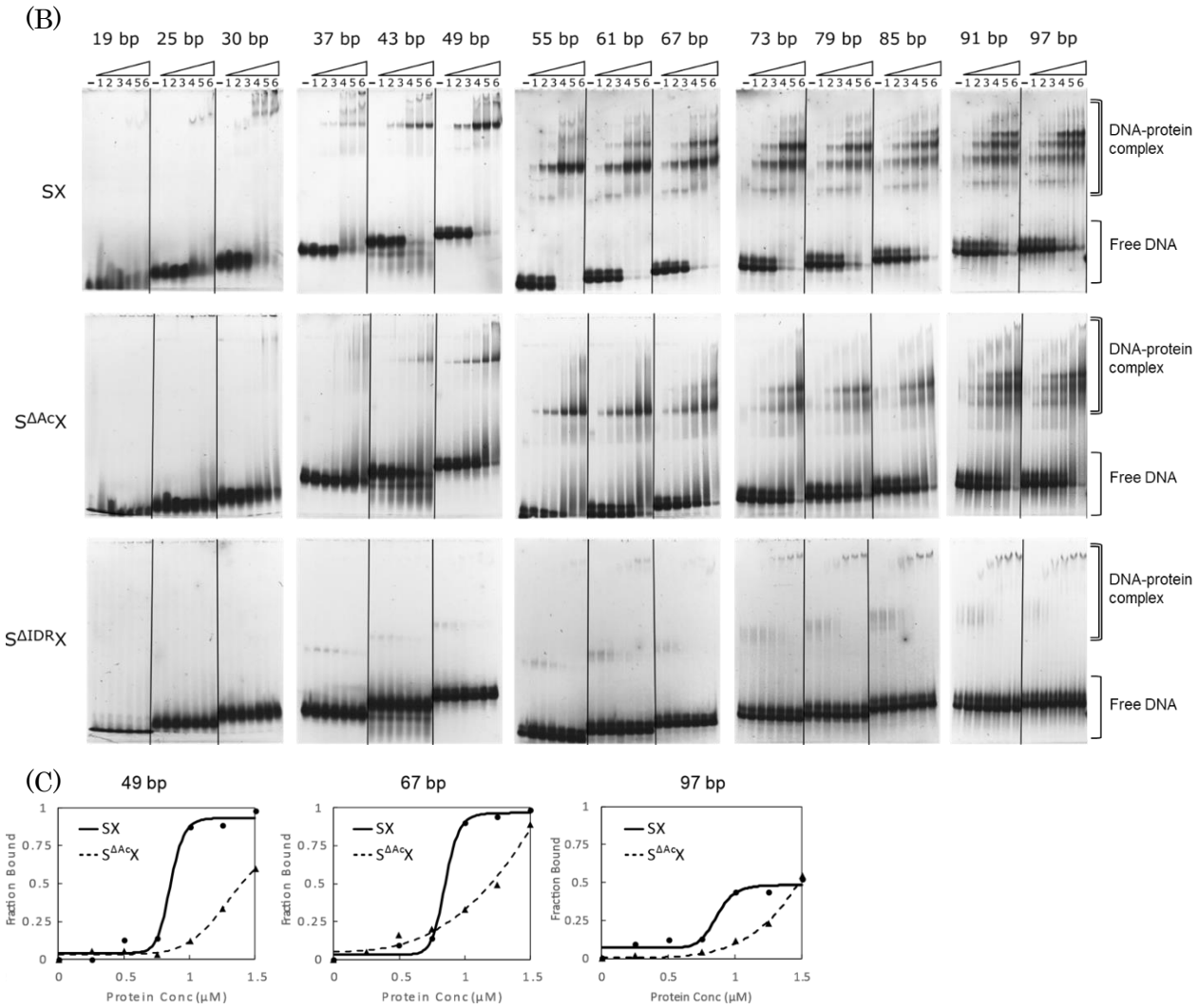


Figure 1. CENP-S C-terminal acidic region determines CENP-SX DNA-binding mode.

(A) Secondary structure prediction plots by Disopred and Sequence analysis of CENP-S C-terminal region. In the plot figure, horizontal and vertical axes show amino acid residue number and prediction value, respectively. The prediction server defines by sequence that if the value is above 0.5, the region is predicted as disordered. In the sequence view, the left column shows Uniprot ID and the first five characters of the organism name. Positively charged amino acids (Lysine, Arginine) and negatively charged amino acids (Glutamate, Aspartate) are colored green and red, respectively. The box region colored light and dark grey in the plot figure is mainly composed by positively and negatively charged residues, respectively. In this study, I prepared three CENP-S expression constructs; full length of CENP-S, CENP-S$^{\Delta Ac}$ (Acidic region deletion), and CENP-S$^{\Delta IDR}$(Intrinsic Disordered Region (IDR) deletion).

(B)Electrophoretic mobility shift assay (EMSA) of CENP-SX, CENP-S$^{\Delta Ac}$X, CENP-S$^{\Delta IDR}$X to a various length of dsDNA and branched structures. 0.25 μM 19-97 bp were mixed with 1, 2, 3, 4, 5, 6-folds equivalents (0.25, 0.5, 0.75, 1.0, 1.25, 1.5 μM) of protein.

(C) Quantitative comparison of CENP-SX and CENP-S$^{\Delta Ac}$X binding 49, 67, and 97 bp. The fraction of bound DNA substrates was plotted against protein concentrations with curve fitting. CENP-SX and CENP-S$^{\Delta Ac}$X were shown as circles with solid lines and triangles with dashed lines, respectively.



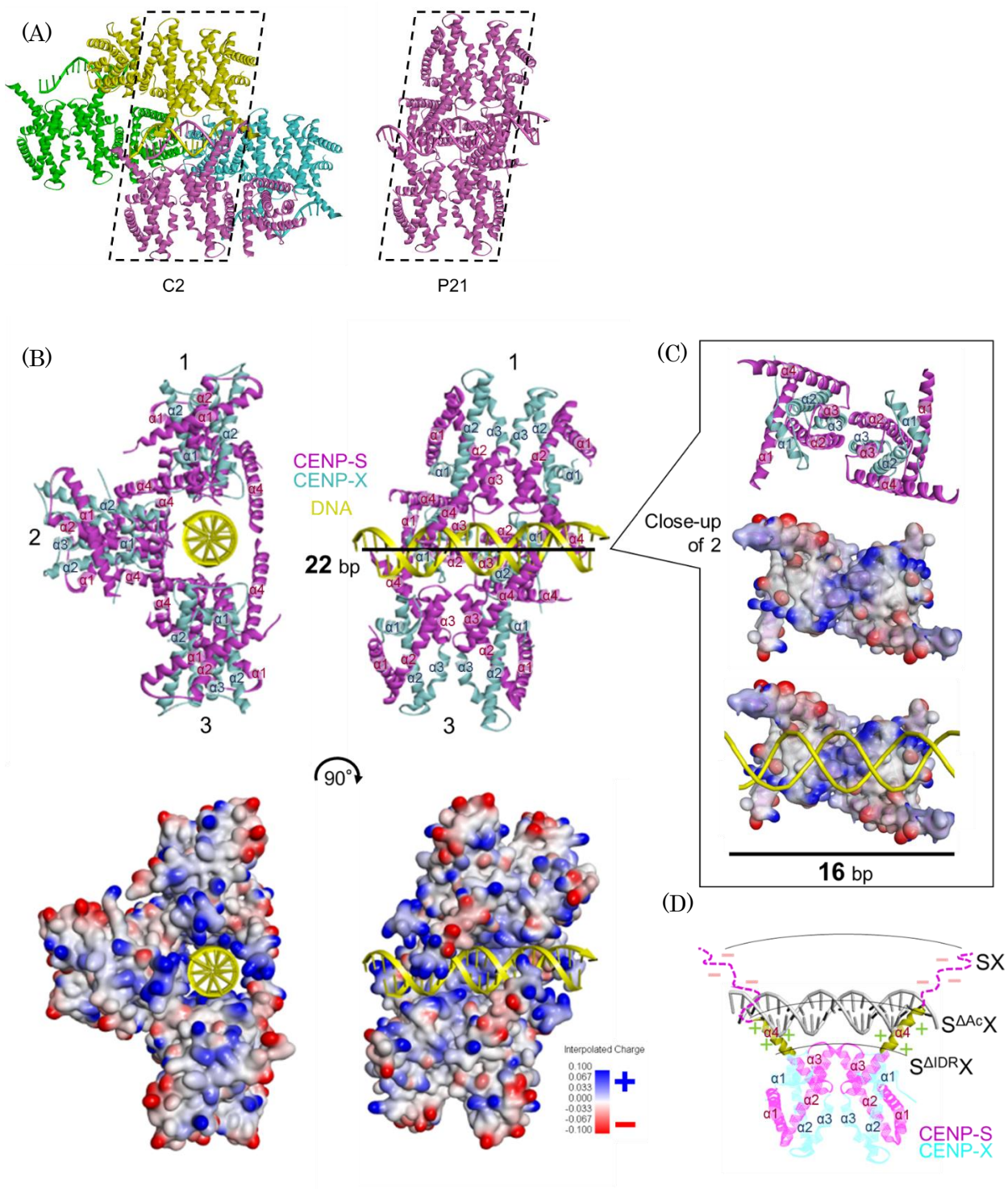


Figure 2. Crystal structure of three CENP-SX heterotetramers bound to single dsDNA.
(A) Compositions for two crystal forms (C2 and $P2_1$) of the CENP-SX-DNA complex are shown. Dashed boxes are identical compositions. Each color is shown as the asymmetric unit.
(B) Crystal structure of CENP-SX-DNA complex. The top panels are ribbon models. Three CENP-SX tetramers are numbered. CENP-S, CENP-X, DNA are colored by cyan, magenta, yellow, respectively. In the bottom panels, CENP-SX tetramers are shown by electrostatic surface; blue and red as positively and negatively charged, respectively. The basic cavity formed within the oligomerized three tetramers accommodates DNA.
(C) CENP-SX top views of ribbon model, electrostatic surface, and its combination with DNA. DNA binding interface by Histone-Fold fits the DNA backbone. A single CENP-SX tetramer binds to 16 bp DNA with support of CENP-S C-terminal α4 extension (see also Figure 2D).
(D) CENP-S α4 positively charged region (colored by yellow) that is predicted as the transitional region from structured to disordered takes α helix structure by DNA binding. Dashed lines show disordered regions unseen in the crystal structure.

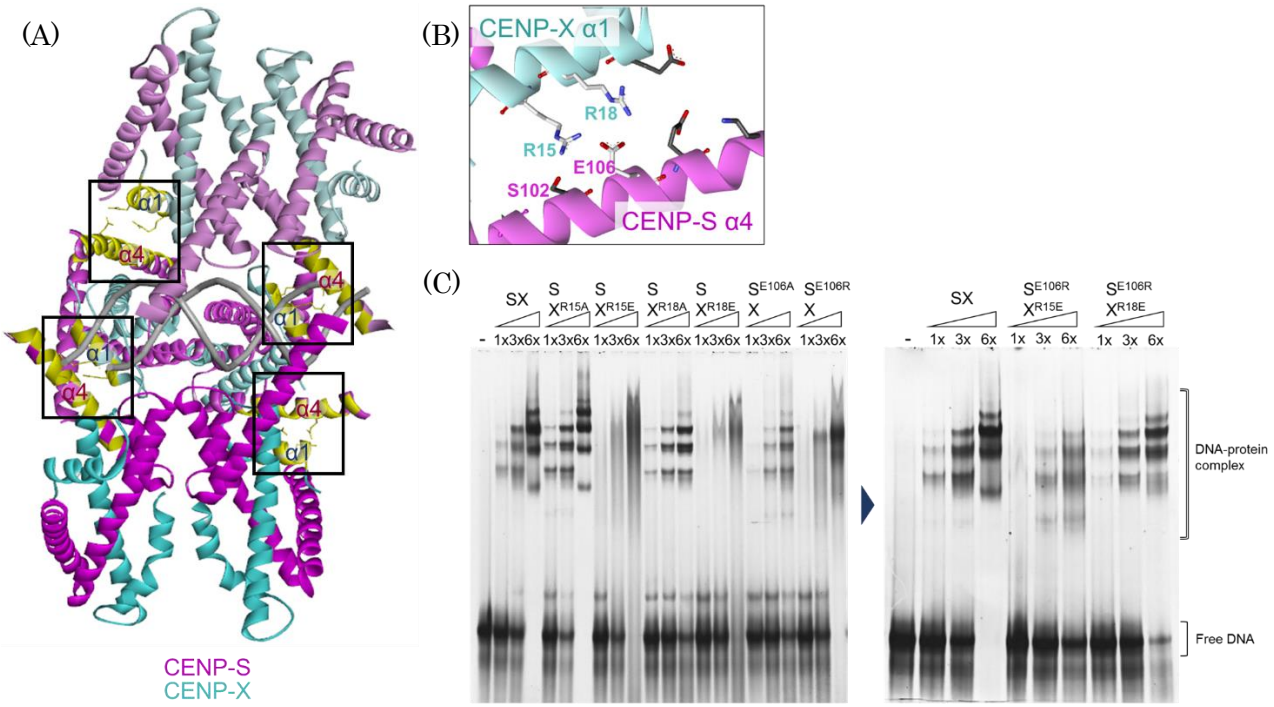


Figure 3. Intermolecular electrostatic interaction is critical for the regular DNA-binding property of the CENP-SX complex.
(A) Intermolecular electrostatic interactions in the CENP-SX-DNA complex are lined square. CENP-S α 4 and CENP-X α 1 are paralleled and colored by yellow. Within the square, E106 (CENP-S) and R15, R18 (CENP-X) are shown by a stick model.
(B) Zoom view of the square from Figure 3A.
(C) Electrophoretic mobility shift assay (EMSA) of 0.25 μM 100 bp from pUC119 DNA with1, 3, 6-folds equivalents (0.25, 0.75, and 1.5 μM) of CENP-SX and its mutants that could compromise the electrostatic interaction by single mutation (left gel) and could restore it by double mutations (right gel).

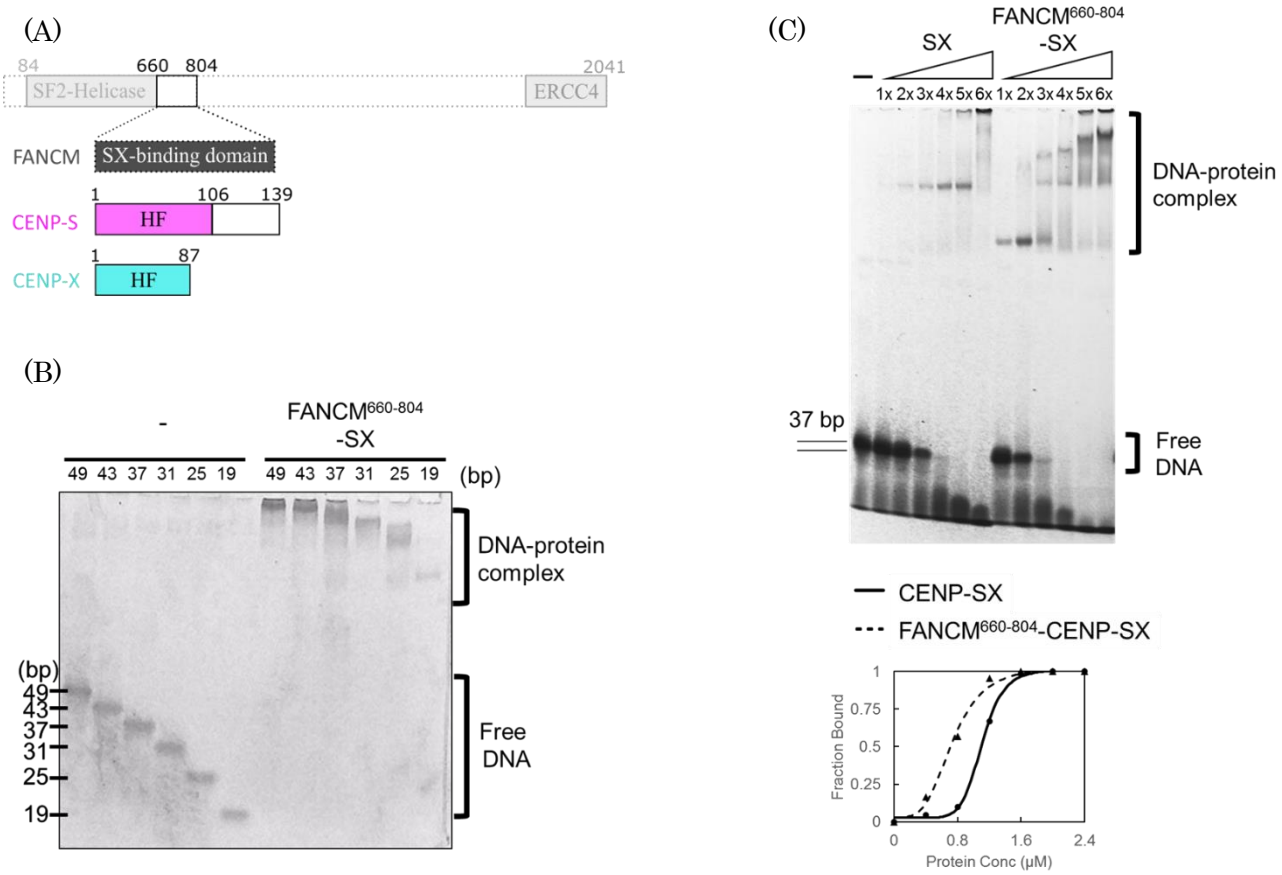


Figure 4. FANCM binding enhances CENP-SX DNA binding.

(A) Schematic domain structure of chicken FANCM, CENP-S, and CENP-X.

FANCM has SF2-Helicase domain at the N-terminus followed by the CENP-SX interaction domain and ERCC4 nuclease-like domain at the C-terminus. CENP-SX-binding indicates an interface with CENP-SX. CENP-SX is composed of CENP-S and CENP-X both containing histone-fold (HF). CENP-S contains an additional intrinsically disordered region.

(B) Electrophoretic mobility shift assay (EMSA) of FANCM$^{660\text{-}804}$-CENP-SX to 19-49 bp DNA. FANCM$^{660\text{-}804}$-CENP-SX-DNA molar ratio was 50 μM:1 μM.

(C) Electrophoretic mobility shift assay (EMSA) and its quantitative comparison of CENP-SX and FANCM$^{660\text{-}804}$-CENP-SX to 37 bp DNA. 0.25 μM DNA were mixed with 1, 2, 3, 4, 5, 6-folds equivalents (0.25, 0.5, 0.75, 1.0, 1.25, 1.5 μM) of protein. The fraction of bound DNA substrates was plotted against protein concentrations with curve fitting. CENP-SX and FANCM$^{660\text{-}804}$-CENP-SX were shown as circles with solid lines and triangles with dashed lines, respectively.

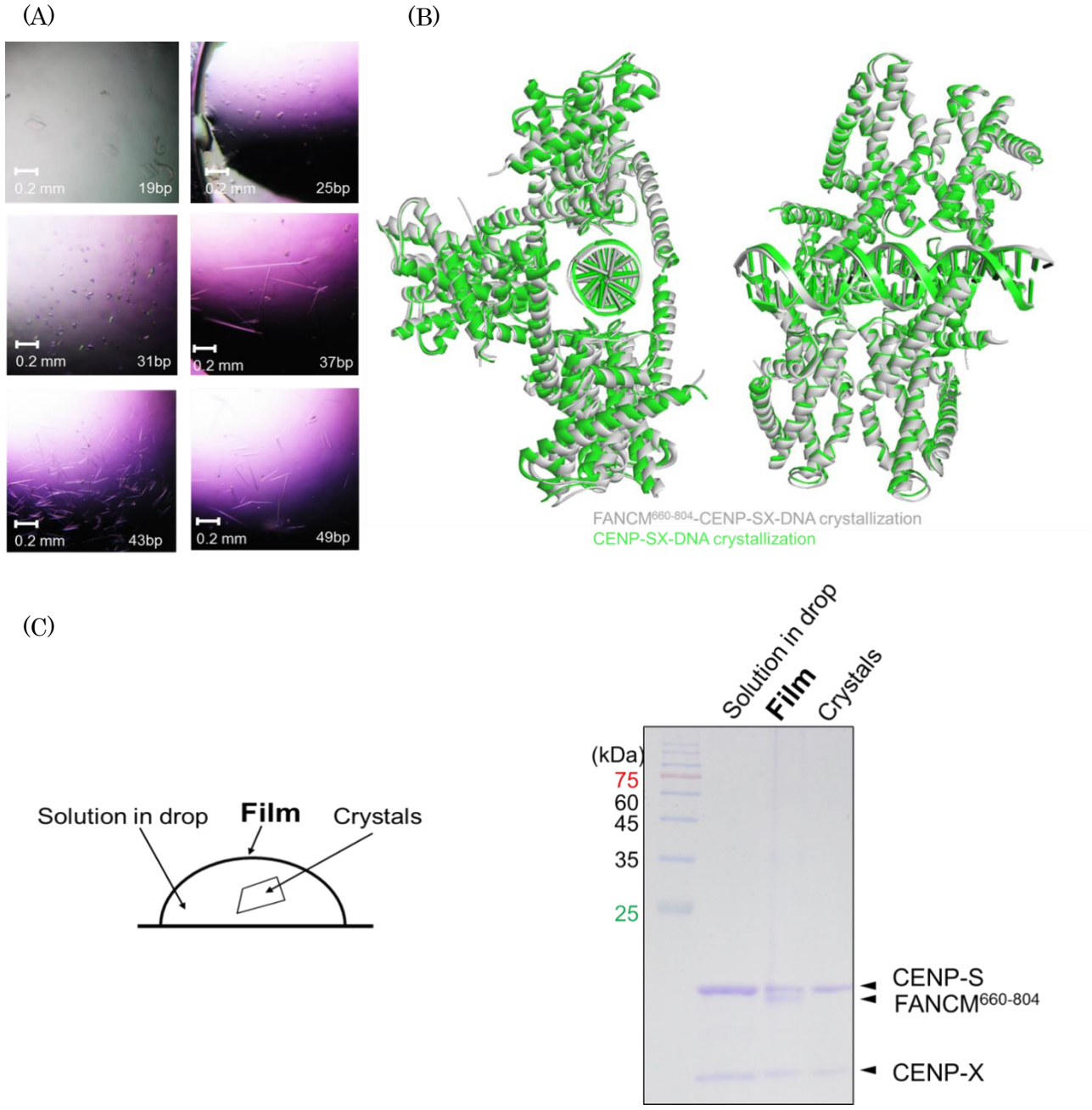


Figure 5. Co-crystallization of FANCM$^{660-804}$-CENP-SX with various length dsDNA resulted in obtaining CENP-SX-DNA complex structure.
(A) Crystal images of FANCM$^{660-804}$-CENP-SX with various length dsDNA.
(B) Superimposition of crystal structures from CENP-SX-DNA crystallization (green) and FANCM$^{660-804}$-CENP-SX-DNA crystallization (grey).
(C) (Left) Schematic drawing of the sitting drop of FANCM$^{660-804}$-CENP-SX-DNA crystallization setup indicating crystals, solution, and the film. Before the crystallization setup, protein-DNA solution (50 μM protein and 55 μM DNA) was prepared. The crystallization drop was formed by mixing of 1 μL of the precipitant solution and 1 μL of protein-DNA solution. A couple of days later, the drop was covered with transparent film. (Right) Analysis of the crystal drop from the condition of FANCM$^{660-804}$-CENP-SX co-crystallization with DNA. Each component was analyzed by 15% SDS-PAGE. The film and crystals were picked up separately, washed within the precipitant solution a few times, and then dissolved in the SDS loading buffer. The gel was stained with Coomassie Brilliant Blue.

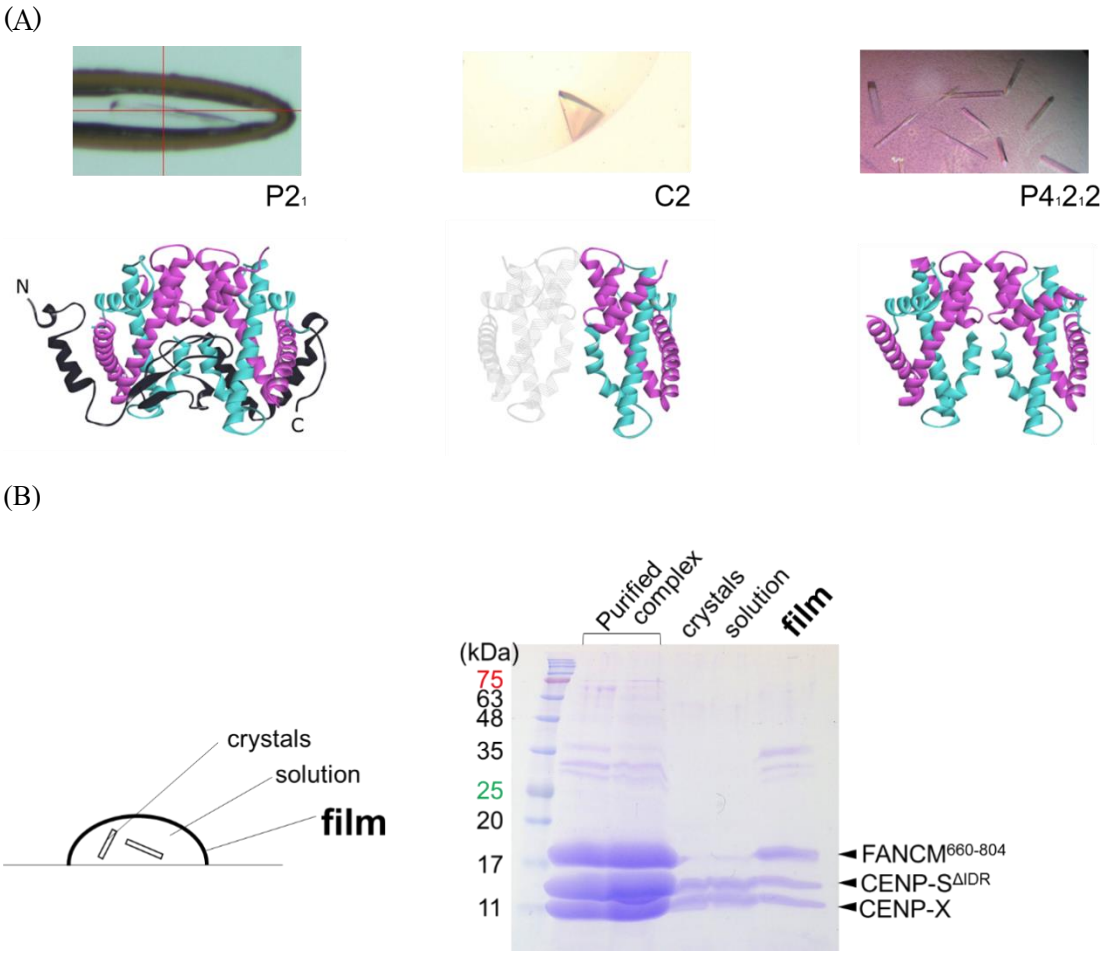


Figure 6. Crystals of FANCM$^{660-804}$-CENP-S$^{\Delta IDR}$X and its analysis.

(A) Photographs of a needle crystal of FANCM$^{660-804}$-CENP-SX (left), a tetrahedral crystal of CENP-S$^{\Delta IDR}$X (middle), and rod-shaped crystals of CENP-S$^{\Delta IDR}$X (right) and their crystal structures. CENP-S$^{\Delta IDR}$, CENP-X, FANCM$^{660-804}$ are colored by cyan, magenta, black, respectively. Crystallographic symmetry heterodimer is colored in the transparent mode in the middle CENP-S$^{\Delta IDR}$X structure. N and C indicate N- and C-terminus of FANCM$^{660-804}$ in the left FANCM-CENP-S$^{\Delta IDR}$X structure.

(B) Analysis of the crystal drop from rod-shaped crystallization condition. (Left) Schematic drawing of the sitting drop crystallization setup indicating crystals, solution, and the film. The crystallization drop was formed by mixing of 1 µL of the precipitant solution and 1 µL of 10mg/mL (176 µM) protein solution. 1-2 weeks later, the drop was covered with transparent film. (Right) Analysis of each component by 15% SDS-PAGE. The film and crystals were picked up separately, washed within the precipitant solution a few times, and then dissolved in the SDS loading buffer. The gel was stained with Coomassie Brilliant Blue.

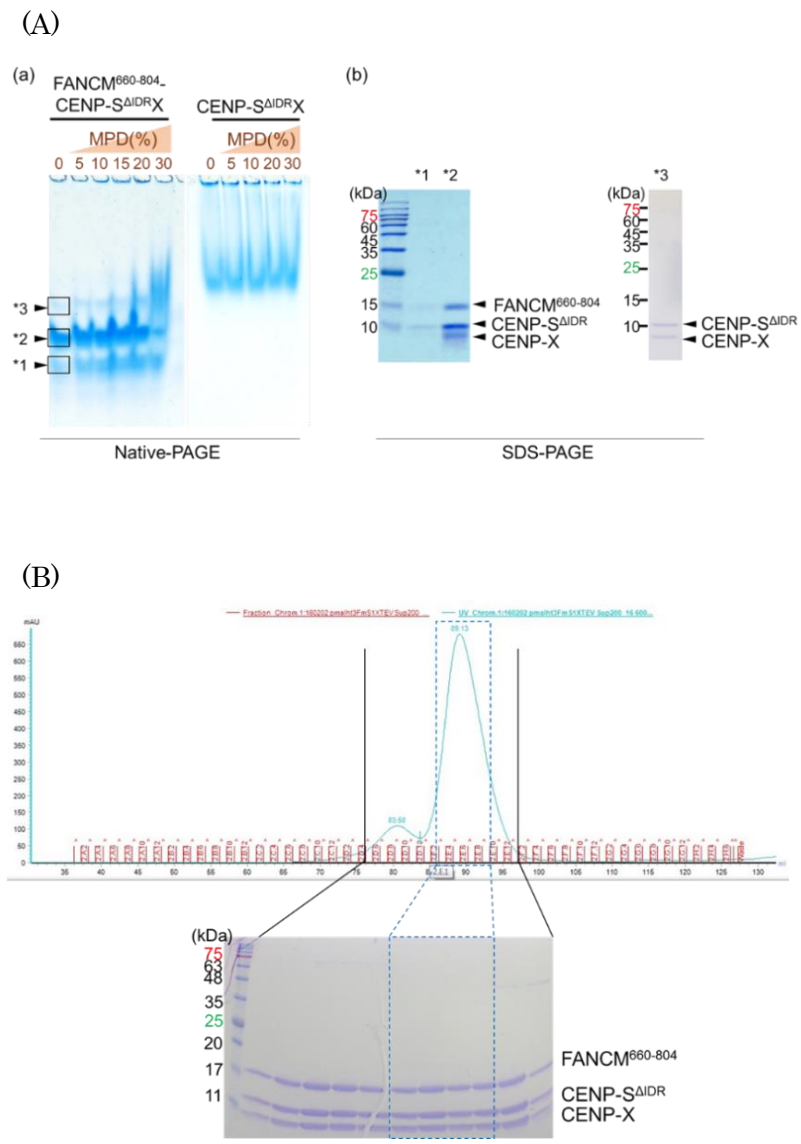


Figure 7. Crystallization condition promotes the release of FANCM$^{660\text{-}804}$ from CENP-S$^{\Delta IDR}$X.

(A) Native-PAGE and SDS-PAGE analysis of FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X and CENP-S$^{\Delta IDR}$X complex in the presence of MPD. (a) FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X (left) and CENP-S$^{\Delta IDR}$X (right) were mixed with different concentrations of MPD and separated on a Native-PAGE. (b) Boxed regions of the Coomassie-stained bands of *1-*3 from (a) were sliced and separated by 15% SDS-PAGE.

(B) Gel filtration profile by Absorbance of 280 nm chromatogram (blue line) and SDS-PAGE analysis of purified FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X complex. The vertical and horizontal axes show Absorbance of 280 nm (mAU) and retention volumes (mL), respectively. Fractions (red short line) surrounded by blue dashed lines were collected and concentrated by ultrafiltration.

(C) Sequence and secondary structure alignment of chicken and human FANCM. Chicken FANCM (GgFANCM) and human FANCM (HsFANCM) were aligned by BLAST pairwise sequence alignment. Numbers indicate amino acid residues in each species. Letters in between the two sequences indicate identical residues and similar residues are marked '+'. Secondary structures of chicken and human FANCM are colored by black and green, respectively. Rectangles, arrows, solid lines represent α-helix, β-sheet, and random coil regions, respectively. Dotted lines are disordered regions. Cysteine residues are highlighted by yellow.

(D) FANCM$^{660\text{-}804}$ is embedded in CENP-SX hydrophobic regions. Ribbon diagrams and Protein surfaces of the chicken FANCM$^{660\text{-}804}$-CENP-S$^{\Delta IDR}$X are displayed using Discovery Studio. Each view of protein surface for hydrophobicity is displayed (brown; hydrophobic, blue; hydrophilic). N and C indicate N- and C-terminus of FANCM$^{660\text{-}804}$, respectively. The hydrophobic area of CENP-SX is covered with FANCM binding. Cys residues are numbered, displayed using CPK style, and colored by yellow. Invisible N-terminal extension of FANCM$^{660\text{-}804}$ in the crystal structure is shown by the broken lines and the cysteine residue within the region is displayed by the yellow circles.

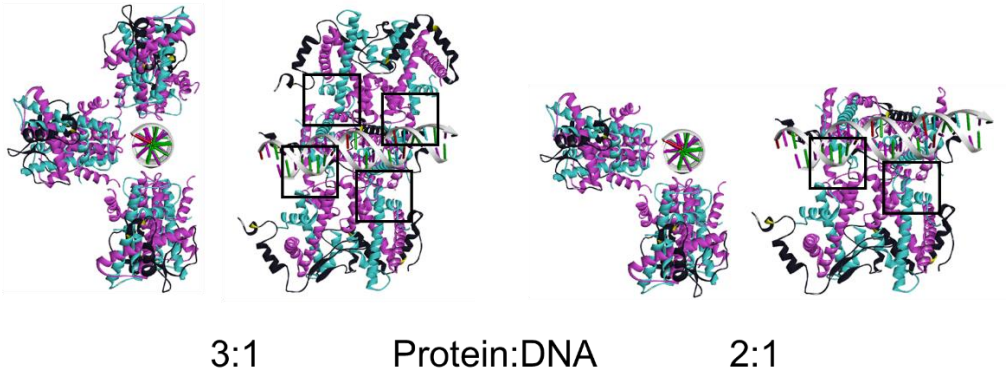


Figure 8. The model of FANCM-CENP-SX chromatin.
CENP-S, CENP-X, FANCM are colored by cyan, magenta, black, respectively. The DNA backbone is shown as a grey arrow. Electrostatic interactions (black squares) between FANCM-CENP-SX molecules could be attained in these models.

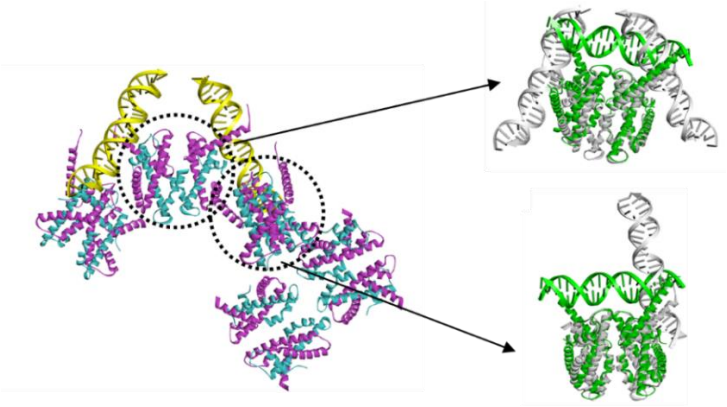


Figure 9. Current CENP-SX DNA binding mode is distinct from that of previous one.
(Left) Human CENP-SX DNA-binding mode from ASU of PDB ID; 4ndy. (Right) Superimposition of my crystal structure (green) with the previously reported crystal structure (grey). CENP-SX tetramer and dimer interface are used for DNA binding in the current and previous structure, respectively.

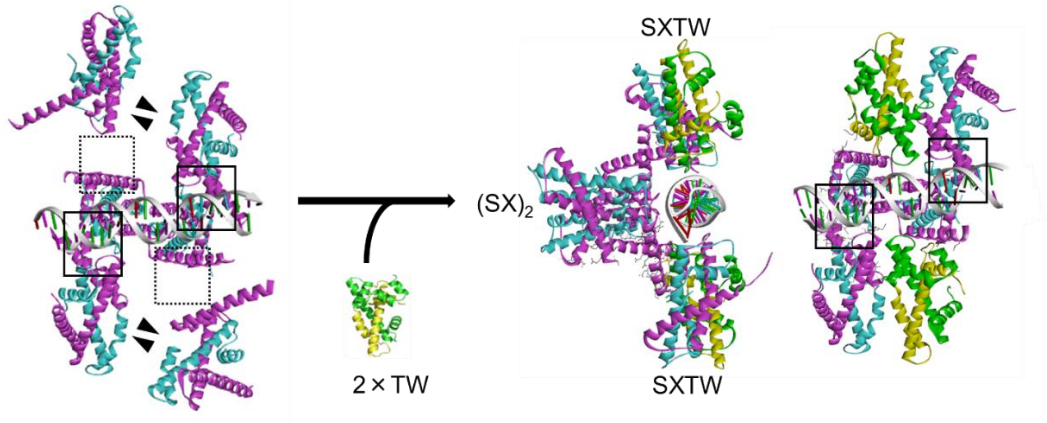


Figure 10. A model of mosaic chromatin composed by CENP-TWSX and CENP-SX.
Black squares indicate electrostatic interactions between CENP-SX molecules. The moiety that is perturbed by CENP-SX dimer dissociation is shown as dashed squares. CENP-S, CENP-X, CENP-T, CENP-W are colored by cyan, magenta, green, yellow, respectively.

Table 1. DNA sequences for EMSA and crystallization.

Description	(bp)		Sequence(5'=>3')
Widom601	19	for	acgtacgcgctgtcccccg
		rev	cgggggacagcgcgtacgt
	25	for	cgcacgtacgcgctgtccccgcgt
		rev	acgcgggggacagcgcgtacgtgcg
	30	for	aaacgcacgtacgcgctgtccccgcgttt
		rev	aaacgcgggggacagcgcgtacgtgcgttt
	37	for	cttaaacgcacgtacgcgctgtccccgcgttttaac
		rev	gttaaaacgcgggggacagcgcgtacgtgcgtttaag
	43	for	ccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgc
		rev	gcggttaaaacgcgggggacagcgcgtacgtgcgtttaagcgg
	49	for	gcaccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgccaa
		rev	ttggcggttaaaacgcgggggacagcgcgtacgtgcgtttaagcggtgc
	55	for	ctagcaccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgccaaggg
		rev	cccttggcggttaaaacgcgggggacagcgcgtacgtgcgtttaagcggtgctag
	61	for	gctctagcaccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgccaaggggat
		rev	atccccttggcggttaaaacgcgggggacagcgcgtacgtgcgtttaagcggtgctagagc
	67	for	acagctctagcaccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgccaaggggattac
		rev	gtaatccccttggcggttaaaacgcgggggacagcgcgtacgtgcgtttaagcggtgctagagctgt
	73	for	tagacagctctagcaccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgccaaggggattactcc
		rev	ggagtaatccccttggcggttaaaacgcgggggacagcgcgtacgtgcgtttaagcggtgctagagctgtcta
	79	for	tcgtagacagctctagcaccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgccaaggggattactcccta
		rev	tagggagtaatccccttggcggttaaaacgcgggggacagcgcgtacgtgcgtttaagcggtgctagagctgtctacga
	85	for	tggtcgtagacagctctagcaccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgccaaggggattactccctagtc
		rev	gactagggagtaatccccttggcggttaaaacgcgggggacagcgcgtacgtgcgtttaagcggtgctagagctgtctacgacca
	91	for	aattggtcgtagacagctctagcaccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgccaaggggattactccctagtctcc
		rev	ggagactagggagtaatccccttggcggttaaaacgcgggggacagcgcgtacgtgcgtttaagcggtgctagagctgtctacgaccaatt
	97	for	ctcaattggtcgtagacagctctagcaccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgccaaggggattactccctagtctccagg
		rev	cctggagactagggagtaatccccttggcggttaaaacgcgggggacagcgcgtacgtgcgtttaagcggtgctagagctgtctacgaccaattgag
pUC119	100	for	caacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaaactggcgaactacttactctagcttcccggcaacaattaa
		rev	ttaattgttgccgggaagctagagtaagtagttcgccagtttaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttg
CENP-SX-DNA crystallization $P2_1$ crystal	30	for	cgcacgtacgcgctgtcccccgcgtttaaa
		rev	aaacgcgggggacagcgcgtacgtgcgttt
CENP-SX-DNA crystallization C2 crystal	31	for	aaacgcacgtacgcgctgtcccccgcgtttt
		rev	aacgcgggggacagcgcgtacgtgcgtttaa

Table 2. Data collection and refinement statistics for CENP-SX-DNA crystallization.

	CENP-SX-DNA	
Wavelength	1.1	1.1
Resolution range	44.38 - 3.2 (3.314 - 3.2)	44.54 - 3.6 (3.729 - 3.6)
Space group	C 1 2 1	P 1 21 1
a, b, c (Å)	127.803 81.527 100.634	100.889 81.616 110.382
α, β, γ (°)	90 123.836 90	90 105.634 90
Total reflections	49154 (5154)	69125 (7187)
Unique reflections	14264 (1425)	20200 (2012)
Multiplicity	3.4 (3.6)	3.4 (3.6)
Completeness (%)	99.32 (99.37)	98.73 (99.45)
Mean I/sigma(I)	24.59 (7.16)	13.04 (4.70)
Wilson B-factor	99.54	110.9
R-merge	0.03158 (0.1642)	0.05473 (0.2357)
R-meas	0.03754 (0.1929)	0.0651 (0.278)
R-pim	0.02007 (0.1006)	0.03486 (0.1461)
CC1/2	0.999 (0.982)	0.999 (0.963)
CC*	1 (0.995)	1 (0.991)
Reflections used in refinement	14248 (1421)	20039 (2004)
Reflections used for R-free	570 (57)	615 (61)
R-work	0.2305 (0.3292)	0.2494 (0.3073)
R-free	0.2614 (0.3602)	0.2928 (0.4205)
CC(work)	0.952 (0.792)	0.963 (0.839)
CC(free)	0.945 (0.868)	0.953 (0.826)
Number of non-hydrogen atoms	4761	9857
macromolecules	4761	9857
Protein residues	548	1117
RMS(bonds)	0.014	0.005
RMS(angles)	1.6	1.07
Ramachandran favored (%)	98.13	98.26
Ramachandran allowed (%)	1.87	1.65
Ramachandran outliers (%)	0	0.09
Rotamer outliers (%)	0	0
Clashscore	28.84	27.6
Average B-factor	121.41	126.34
macromolecules	121.41	126.34
Number of TLS groups	17	38

Statistics for the highest-resolution shell are shown in parentheses.

Table 3. Data collection and refinement statistics for FANCM$^{660-804}$-CENP-S$^{\Delta IDR}$X crystallization.

	CENP-S$^{\Delta IDR}$X (from tetrahedral crystal)	CENP-S$^{\Delta IDR}$X (from rod-shaped crystal)	FANCM$^{660-804}$-CENP-S$^{\Delta IDR}$X
PDB entry	**7da0**	**7da1**	**7da2**
Wavelength	0.9	1.1	0.98
Resolution range	34.01 - 1.25 (1.295 - 1.25)	36.5 - 2.01 (2.082 - 2.01)	45.17 - 2.79 (2.89 - 2.79)
Space group	C 1 2 1	P 41 21 2	P 21 21 21
a, b, c (Å)	50.566, 69.071, 48.878	59.432, 59.432, 220.949	71.249, 78.469, 87.468
α, β, γ (°)	90, 103.666, 90	90, 90, 90	90, 90, 90
Total reflections	160822	347250	76771
Unique reflections	43638 (4340)	27443 (2661)	12641 (1232)
Multiplicity	3.7	12.6	6
Completeness (%)	96.74 (96.21)	99.85 (99.81)	99.32 (99.84)
Mean I/sigma(I)	36.64 (2.17)	22.3 (3.0)	32.0 (4.32)
Wilson B-factor	15.4	24.66	55.02
R-merge	0.056 (0.851)	0.103 (0.772)	0.133 (0.788)
R-meas	0.066 (0.999)	0.108 (0.802)	0.145 (0.873)
R-pim	0.034 (0.52)	0.030 (0.217)	0.059 (0.37)
CC1/2[a]	NA (0.407)	NA (0.93)	NA (0.427)
CC*[a]	NA (0.761)	NA (0.982)	NA (0.774)
Reflections used in refinement	43628 (4337)	27434 (2662)	12622 (1232)
Reflections used for R-free	2000 (199)	1988 (193)	632 (63)
R-work	0.1927 (0.2966)	0.1968 (0.2075)	0.2161 (0.2829)
R-free	0.2124 (0.2958)	0.2244 (0.2563)	0.2616 (0.3522)
Number of non-hydrogen atoms	1619	2878	3858
macromolecules	1358	2704	3811
solvent	249	174	47
Protein residues	169	338	469
RMS(bonds)	0.009	0.007	0.003
RMS(angles)	0.98	0.91	0.53
Ramachandran favored (%)	100	99.7	94.77
Ramachandran allowed (%)	0	0.3	4.36
Ramachandran outliers (%)	0	0	0.87
Rotamer outliers (%)	2.11	0	3.43
Clashscore	5.44	5.82	7.85
Average B-factor	24.88	31.39	66.05
macromolecules	22.73	31.03	66.02
solvent	36.02	36.91	68.26
Number of TLS groups	6	15	20

Statistics for the highest-resolution shell are shown in parentheses.

[a]Average CC1/2 and CC* values were not reported by version of HKL2000/Scalepack.

Acknowledgments

I am deeply grateful to Dr. Nishino for his extensive assistance and advice. I appreciate the help of our laboratory members. I'd like to thank Dr. Miyazaki, Dr. Toshima, Dr. Arimura, and Dr. Shimizu for helpful discussions and comments to this article.

References

1. M. H. Hauer, S. M. Gasser, Chromatin and nucleosome dynamics in DNA damage and repair. *Genes Dev.* **31**, 2204–2221 (2017).
2. M. Hara, T. Fukagawa, Dynamics of kinetochore structure and its regulations during mitotic progression. *Cell. Mol. Life Sci.* **77**, 2981–2995 (2020).
3. M. Perpelescu, T. Fukagawa, The ABCs of CENPs. *Chromosoma*. **120**, 425–446 (2011).
4. T. Nishino, K. Takeuchi, K. E. Gascoigne, A. Suzuki, T. Hori, T. Oyama, K. Morikawa, I. M. Cheeseman, T. Fukagawa, CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. *Cell*. **148**, 487–501 (2012).
5. M. Amano, A. Suzuki, T. Hori, C. Backer, K. Okawa, I. M. Cheeseman, T. Fukagawa, The CENP-S complex is essential for the stable assembly of outer kinetochore structure. *J. Cell Biol.* **186**, 173–182 (2009).
6. T. R. Singh, D. Saro, A. M. Ali, X.-F. Zheng, C. Du, M. W. Killen, A. Sachpatzidis, K. Wahengbam, A. J. Pierce, Y. Xiong, P. Sung, A. R. Meetei, MHF1-MHF2, a Histone-Fold-Containing Protein Complex, Participates in the Fanconi Anemia Pathway via FANCM. *Mol. Cell*. **37**, 879–886 (2010).
7. Z. Yan, M. Delannoy, C. Ling, D. Daee, F. Osman, P. A. Muniandy, X. Shen, A. B. Oostra, H. Du, J. Steltenpool, T. Lin, B. Schuster, C. Décaillet, A. Stasiak, A. Z. Stasiak, S. Stone, M. E. Hoatlin, D. Schindler, C. L. Woodcock, H. Joenje, R. Sen, J. P. de Winter, L. Li, M. M. Seidman, M. C. Whitby, K. Myung, A. Constantinou, W. Wang, A Histone-Fold Complex and FANCM Form a Conserved DNA-Remodeling Complex to Maintain Genome Stability. *Mol. Cell*. **37**, 865–878 (2010).
8. F. Zafar, A. K. Okita, A. T. Onaka, J. Su, Y. Katahira, J. I. Nakayama, T. S. Takahashi, H. Masukata, T. Nakagawa, Regulation of mitotic recombination between DNA repeats in centromeres. *Nucleic Acids Res.* **45**, 11222–11235 (2017).
9. C. Girard, W. Crismani, N. Froger, J. Mazel, A. Lemhemdi, C. Horlow, R. Mercier, FANCM-associated proteins MHF1 and MHF2, but not the other Fanconi anemia factors, limit meiotic crossovers. *Nucleic Acids Res.* **42**, 9087–9095 (2014).
10. S. Bhattacharjee, F. Osman, L. Feeney, A. Lorenz, C. Bryer, M. C. Whitby, MHF1-2/CENP-S-X performs distinct roles in centromere metabolism and genetic recombination. *Open Biol.* **3**, 130102 (2013).
11. J. Huang, S. Liu, M. A. Bellani, A. K. Thazhathveetil, C. Ling, J. P. DeWinter, Y. Wang, W. Wang, M. M. Seidman, The DNA Translocase FANCM/MHF Promotes Replication Traverse of DNA Interstrand Crosslinks. *Mol. Cell*. **52**, 434–446 (2013).
12. G. Milletti, L. Strocchio, D. Pagliara, K. Girardi, R. Carta, A. Mastronuzzi, F. Locatelli, F. Nazio, Canonical and noncanonical roles of fanconi anemia proteins: Implications in cancer predisposition. *Cancers (Basel)*. **12**, 1–23 (2020).
13. A. J. Deans, S. C. West, FANCM Connects the Genome Instability Disorders Bloom's Syndrome

and Fanconi Anemia. *Mol. Cell*. **36**, 943–953 (2009).

14. S. Nandi, M. C. Whitby, The ATPase activity of Fml1 is essential for its roles in homologous recombination and DNA repair. *Nucleic Acids Res.* **40**, 9584–9595 (2012).
15. N. E. Romero, S. W. Matson, J. Sekelsky, Biochemical activities and genetic functions of the Drosophila melanogaster fancm helicase in DNA repair. *Genetics*. **204**, 531–541 (2016).
16. W. Crismani, C. Girard, N. Froger, M. Pradillo, J. L. Santos, L. Chelysheva, G. P. Copenhaver, C. Horlow, R. Mercier, FANCM limits meiotic crossovers. *Science*. **336**, 1588–1590 (2012).
17. X. F. Zheng, R. Prakash, D. Saro, S. Longerich, H. Niu, P. Sung, Processing of DNA structures via DNA unwinding and branch migration by the S. cerevisiae Mph1 protein. *DNA Repair (Amst).* **10**, 1034–1043 (2011).
18. W. Sun, S. Nandi, F. Osman, J. S. Ahn, J. Jakovleska, A. Lorenz, M. C. Whitby, The FANCM Ortholog Fml1 Promotes Recombination at Stalled Replication Forks and Limits Crossing Over during DNA Double-Strand Break Repair. *Mol. Cell*. **32**, 118–128 (2008).
19. K. Gari, C. Décaillet, A. Z. Stasiak, A. Stasiak, A. Constantinou, The Fanconi Anemia Protein FANCM Can Promote Branch Migration of Holliday Junctions and Replication Forks. *Mol. Cell*. **29**, 141–148 (2008).
20. K. Gari, C. Décaillet, M. Delannoy, L. Wu, A. Constantinou, Remodeling of DNA replication structures by the branch point translocase FANCM. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16107–16112 (2008).
21. I. V. Rosado, W. Niedzwiedz, A. F. Alpi, K. J. Patel, The Walker B motif in avian FANCM is required to limit sister chromatid exchanges but is dispensible for DNA crosslink repair. *Nucleic Acids Res.* **37**, 4360–4370 (2009).
22. X. Xue, P. Sung, X. Zhao, Functions and regulation of the multitasking FANCM family of DNA motor proteins. *Genes Dev.* **29**, 1777–1788 (2015).
23. A. R. Meetei, A. L. Medhurst, C. Ling, Y. Xue, T. R. Singh, P. Bier, J. Steltenpool, S. Stone, I. Dokal, C. G. Mathew, M. Hoatlin, H. Joenje, J. P. De Winter, W. Wang, A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat. Genet.* **37**, 958–963 (2005).
24. M. C. Whitby, The FANCM family of DNA helicases/translocases. *DNA Repair (Amst).* **9**, 224–236 (2010).
25. Y. Tao, C. Jin, X. Li, S. Qi, L. Chu, L. Niu, X. Yao, M. Teng, The structure of the FANCMMHF complex reveals physical features for functional assembly. *Nat. Commun.* **3**, 1–12 (2012).
26. D. Fox, Z. Yan, C. Ling, Y. Zhao, D. Y. Lee, T. Fukagawa, W. Yang, W. Wang, The histone-fold complex MHF is remodeled by FANCM to recognize branched DNA and protect genome stability. *Cell Res.* **24**, 560–575 (2014).
27. Q. Zhao, D. Saro, A. Sachpatzidis, T. R. Singh, D. Schlingman, X. F. Zheng, A. MacK, M. S. Tsai, S. Mochrie, L. Regan, A. R. Meetei, P. Sung, Y. Xiong, The MHF complex senses branched DNA by binding a pair of crossover DNA duplexes. *Nat. Commun.* **5**, 1–12 (2014).

28. K. Anand, D. Pal, R. Hilgenfeld, An overview on 2-methyl-2,4-pentanediol in crystallization and in crystals of biological macromolecules. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58**, 1722–1728 (2002).
29. Y. Takizawa, C. H. Ho, H. Tachiwana, H. Matsunami, W. Kobayashi, M. Suzuki, Y. Arimura, T. Hori, T. Fukagawa, M. D. Ohi, M. Wolf, H. Kurumizaka, Cryo-EM Structures of Centromeric Tri-nucleosomes Containing a Central CENP-A Nucleosome. *Structure*. **28**, 44-53.e4 (2020).
30. K. Takeuchi, T. Nishino, K. Mayanagi, N. Horikoshi, A. Osakabe, H. Tachiwana, T. Hori, H. Kurumizaka, T. Fukagawa, The centromeric nucleosome-like CENP-T-W-S-X complex induces positive supercoils into DNA. *Nucleic Acids Res.* **42**, 1644–1655 (2014).
31. G. K. Hom, Dioxane contributes to the altered conformation and oligomerization state of a designed engrailed homeodomain variant. *Protein Sci.* **14**, 1115–1119 (2005).
32. D. T. Jones, D. Cozzetto, DISOPRED3: Precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*. **31**, 857–863 (2015).
33. D. W. A. Buchan, D. T. Jones, The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res.* **47**, W402–W407 (2019).
34. D. T. Jones, Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195–202 (1999).
35. W. Kabsch, XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–32 (2010).
36. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
37. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).