Cell Research (2014) 24:560-575.

www.nature.com/cr

ORIGINAL ARTICLE

The histone-fold complex MHF is remodeled by FANCM to recognize branched DNA and protect genome stability

David Fox 3rd[1, *], Zhijiang Yan[1, *], Chen Ling[1], Ye Zhao[2], Duck-Yeon Lee[3], Tatsuo Fukagawa[4], Wei Yang[2], Weidong Wang[1]

[1]Laboratory of Genetics, National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA; [2]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Health, Bethesda, MD 20892, USA; [3]Biochemistry Core Facility, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA; [4]Department of Molecular Genetics, National Institute of Genetics and the Graduate University for Advanced Studies, Mishima 411-8540, Japan

Histone-fold proteins typically assemble in multiprotein complexes to bind duplex DNA. However, one histone-fold complex, MHF, associates with Fanconi anemia (FA) protein FANCM to form a branched DNA remodeling complex that senses and repairs stalled replication forks and activates FA DNA damage response network. How the FANCM-MHF complex recognizes branched DNA is unclear. Here, we solved the crystal structure of MHF and its complex with the MHF-interaction domain (referred to as MID) of FANCM, and performed structure-guided mutagenesis. We found that the MID-MHF complex consists of one histone H3-H4-like MHF heterotetramer wrapped by a single polypeptide of MID. We identified a zinc atom-liganding structure at the central interface between MID and MHF that is critical for stabilization of the complex. Notably, the DNA-binding surface of MHF was altered by MID in both electrostatic charges and allosteric conformation. This leads to a switch in the DNA-binding preference — from duplex DNA by MHF alone, to branched DNA by the MID-MHF complex. Mutations that disrupt either the composite DNA-binding surface or the protein-protein interface of the MID-MHF complex impaired activation of the FA network and genome stability. Our data provide the structural basis of how FANCM and MHF work together to recognize branched DNA, and suggest a novel mechanism by which histone-fold complexes can be remodeled by their partners to bind special DNA structures generated during DNA metabolism.

***Keywords*:** FANCM; Fanconi anemia; MHF; histone-fold; genome stability; branched DNA

Cell Research (2014) **24**:560-575. doi:10.1038/cr.2014.42; published online 4 April 2014

Introduction

Fanconi anemia (FA) is a rare genetic syndrome featuring genomic instability, bone marrow failure and cancer predisposition [1-3]. The hallmark of FA cells is hypersensitivity to drugs that induce DNA interstrand crosslinks (ICLs), which block DNA replication and require resolution prior to genome duplication. FA is caused by mutations in any of 16 different FANC genes (FANCA to FANCQ) [4, 5]. The products of these genes act in the FA DNA damage response network, called FA network, which functions to repair ICLs and restore blocked replication forks [1-3].

FANCM is a crucial FANC protein [6, 7] and plays several roles in the FA network. For example, FANCM catalyzes ATP-dependent remodeling of branched DNA structures produced during replication and repair, including replication forks and Holliday junctions (HJs) [8-11]. This remodeling activity is required for FANCM to recover stalled replication forks [12-15], promote cellular resistance to ICLs [10, 16, 17], and suppress sister chromatid exchanges (SCEs) [18]. In addition, FANCM preferentially binds replication forks and HJs through its helicase domain [9, 10]. This ATP-independent binding activity is required for FANCM to recruit the FA core complex to chromatin, enabling the ubiquitin ligase subunit of the complex (FANCL) to monoubiquitinate

*These two authors contributed equally to this work.
Correspondence: Weidong Wang
Tel: +1-410-558-8334; Fax: +1-410-558-8331
E-mail: wangw@grc.nia.nih.gov
Received 6 November 2013; revised 10 February 2014; accepted 24 February 2014; published online 4 April 2014

FANCD2 and FANCI [10]. Both activities of FANCM are stimulated by its interacting partner, MHF, which consists of a heterodimer of histone-fold proteins, MHF1 and MHF2 [14, 19]. FANCM and MHF form a complex, FANCM-MHF, which is conserved from yeast to human and is essential for genome protection in both species [14]. Only a small fraction of FANCM-MHF associates with and recruits the FA core complex to chromatin, leading to monoubiquitination of FANCD2 and FANCI [14].

Structural studies have shown that the MHF complex is a heterotetramer that comprises two MHF1-MHF2 heterodimers, and is structurally similar to histone H3-H4 heterotetramer [20-22]. Like nucleosomes and other histone-fold complexes, MHF preferentially binds double-strand DNA (dsDNA) [14]. In fact, MHF1 and MHF2 have been identified as the centromeric proteins CENP-S and CENP-X, respectively, both of which promote kinetochore assembly [23-25]. In this case, they associate with another histone-fold heterodimer, CENP-T/CENP-W, to form a mixed histone-fold heterotetramer, CENP-T-W-S-X, which binds and supercoils dsDNA in a manner similar to nucleosomes [20]. It is not surprising that MHF mediates the dsDNA binding of the CENP-T-W-S-X complex because its preferred substrate is dsDNA. To explain how MHF facilitates the binding and remodeling of branched DNA by FANCM, we proposed previously that the two proteins cooperatively bind distinct parts of a branched DNA molecule: FANCM at the branch point, and MHF at the surrounding duplex regions [14].

An earlier study has reported crystal structure of a complex between the MHF-interacting domain of FANCM (residues 661-800; abbreviated as MID) and MHF at 2.64 Å resolution. This study revealed a pentameric structure containing one MID polypeptide interacting with the MHF1-MHF2 heterotetramer, and identified a novel composite DNA-binding surface [22]. However, the relevance of these structures to the function of FANCM-MHF in the FA network and genome stability has not been addressed. Here, we describe higher resolution crystal structures of both MHF and MID-MHF complexes, which were solved to 2.1 Å and 2.0 Å resolution, respectively. We identified a novel zinc atom-liganding structure required for stabilizing the FANCM-MHF complex. Importantly, we observed asymmetric remodeling of the DNA-binding surface of MHF by MID, which leads to a switch in the DNA-binding preference: from dsDNA by MHF to branched DNA by MID-MHF complex. We demonstrated that both the composite DNA-binding surface and the protein-protein interface of the FANCM-MHF complex are important for activation of the FA network and protection of genome stability.

Results

Identification of a zinc atom-liganding structure at the central interface of the MID-MHF pentamer

To understand the mechanism of action of the MID-MHF complex, we solved the crystal structures of the MHF and MID-MHF complexes (Figure 1A, Supplementary information, Figure S1A-S1D and Table S1). Consistent with previous structural studies [20-22], MHF is a histone-fold heterotetramer containing two MHF1-MHF2 heterodimers, similar to that of the histone H3-H4 heterotetramer (Supplementary information, Figure S1D); and the MID-MHF complex is a pentamer composed of a single MID polypeptide (residues 676-791) that wraps in a circuitous fashion around the MHF tetramer (Figure 1A). Interestingly, a single metal atom liganding site was observed at the interface between the C-terminus of MHF2 (chain-D) and MID (L4-α3-helix), with each protein contributing half of the total bidentate-binding motif with tetrahedral coordination geometry (Figure 1B). The metal atom was subsequently identified by atomic absorption spectroscopy as a single zinc atom present in only MID-MHF complex, but not in MHF alone (Figure 1C).

It should be noted that atomic absorption spectroscopy cannot rule out the possibility that other metal ions might substitute Zn^{2+} at this site. However, we favor the notion that this metal ion is Zn^{2+} because: (1) Zn^{2+} is much more prevalent than other metal ions with two positive charges in stabilizing structures of protein and protein complexes; (2) the observed crystal structure of the metal ion with its coordinating residues is consistent with tetrahedral coordination geometry for the structural Zn^{2+} ion; and (3) mutations of two commonly used Zn coordination residues, Cys and His, disrupt MID-MHF association and function (see below).

It should also be mentioned that Tao *et al.* [22] have reported a similar pentameric structure and hydrophobic interface for the MID-MHF complex. However, the earlier structure has lower resolution than the structure presented here (2.64 Å vs 2.0 Å), and it is missing the structural zinc atom and its liganding residues that mediate critical interactions between MID and MHF (see below). Moreover, the functional relevance of the observed protein interfaces was not addressed.

The zinc atom-liganding structure and hydrophobic interfaces within the central motifs of MID are essential for stabilization of the MID-MHF pentamer

We selectively mutated the most conserved residues in the MID-MHF pentamer (Supplementary information, Figures S2, S3 and Table S2), which likely play essen-

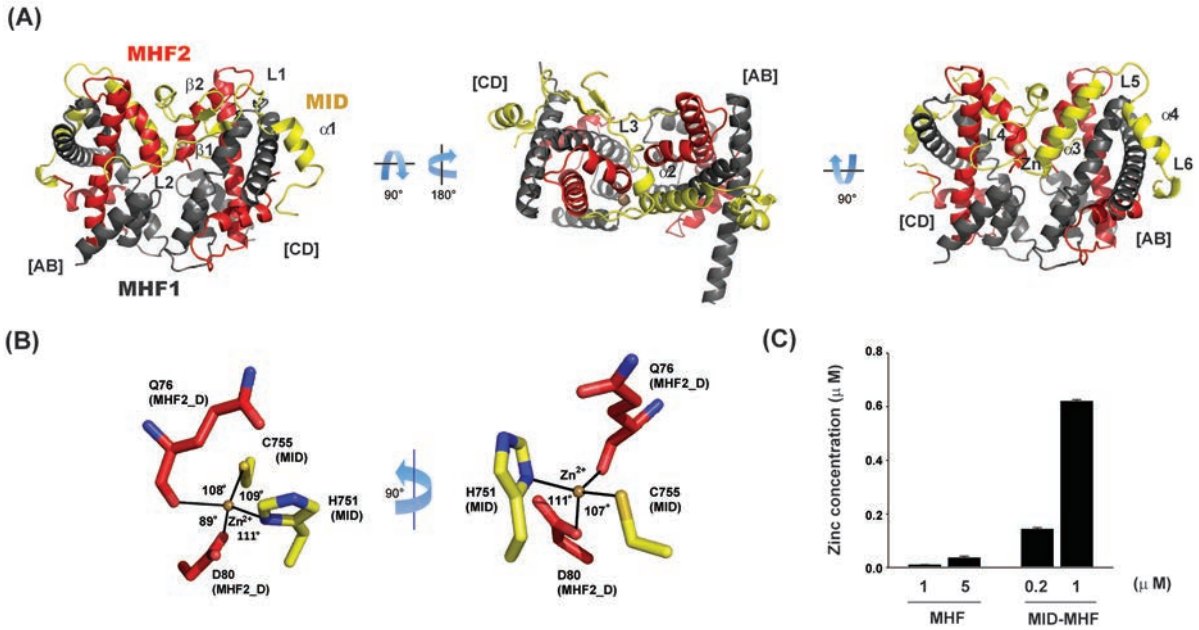

Figure 1 Identification of a zinc atom-liganding structure at the central interface of the MID-MHF pentamer. **(A)** Crystal structure of the MID-MHF complex with three orthogonal orientations (MHF1, dark gray; MHF2, red and MID, yellow as indicated). The individual MHF1/MHF2 heterodimers are labeled [AB] or [CD] to denote the chains relating to the deposited crystal structure (PDB ID: 4E44), and the secondary structure is labeled accordingly (MHF1, black, α1- α4; MHF2, red, α1- α3). MID secondary structure is labeled according to the deposited crystal structure (PDB ID: 4E45). The structural zinc (Zn) atom is shown as a brown sphere. **(B)** Coordination of the structural zinc atom with MHF2 chain D (red) and MID (yellow) residues coming together to form the bidentate zinc-chelating motif. Coordination includes the MHF2 Q76 carbonyl oxygen and the carboxylic acid side-chain of D80 and the MID H751 imidazole and C755 thiol side-chains. **(C)** Atomic absorption analyses showing the presence of Zn in the MID-MHF complex, but not in the MHF complex. Measured concentrations are the average of triplicate analyses.

tial roles in stabilizing the complex. A majority of such residues in MID are located in its central motifs from L3 to α3 (Supplementary information, Figure S3), including those that transit between α3 helices of MHF2 (Figure 1A). Notably, the strictly conserved residues within this region are either hydrophobic (including L733, V749, F758 and M762) or involved in Zn coordination (H751), with none being hydrophilic (Supplementary information, Figure S3). This implies that the interactions mediated by hydrophobic and Zn coordinating residues are likely to be more critical for MID-MHF assembly than those through hydrophilic ones.

To investigate this prediction, we made three groups of point mutants. The first group has mutated residues H751 and C755, which coordinate the zinc atom along with MHF2 Q76 carbonyl oxygen and D80 side-chain of the α3 D-chain (Figure 1B). The second group includes residues that engage in hydrophobic interactions with MHF, such as V749 single and F758/M762 double mutant (Figure 2A and 2B). The third group has a mutated residue R754, which is the most conserved hydrophilic residue within this region (with only one variation to Leu) (Supplementary information, Figures S3 and S4A). Only the first two groups of mutants, but not the third, co-immunoprecipitated with reduced levels of MHF1 and MHF2 (Figure 2C and 2D). The data are consistent with the evolution-based prediction that interactions through Zn coordinating and hydrophobic residues, but not hydrophilic residues, are critical for normal FANCM-MHF complex assembly. We also made the V749/H751 double mutant to disrupt both hydrophobic and Zn-mediated interactions. Whereas each single mutant still had residual association with MHF (Figure 2C), the double mutant was completely deficient (Figure 2D). These data further support the notion that normal FANCM-MHF complex assembly requires both Zn coordination and hydrophobic interactions mediated by the central motifs of MID.

We further analyzed the association between MID variants and MHF by Superose 6 gel-filtration chromatography using whole-cell lysates of HEK293 cells transfected with several Flag-MID variants, one with normal (L733G) and two with defective association

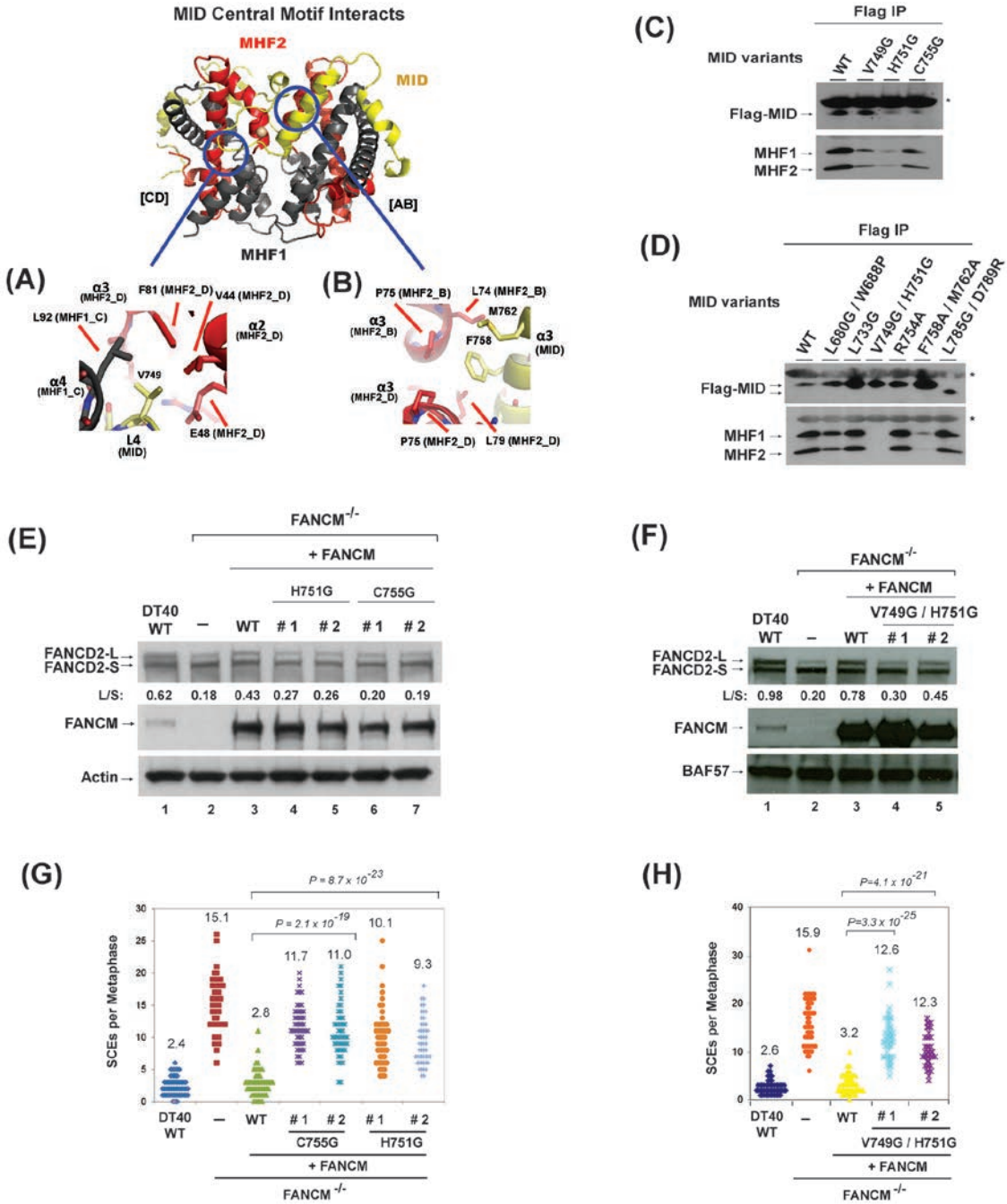

Figure 2 Identification of the residues in MID that are essential for stabilization of the MID-MHF complex, normal activation of the FA network and suppression of SCEs. **(A**, **B)** Key MID and MHF interfaces are shown in detail. MID central motif residues: **(A)** L4 (MID-V749), **(B)** α3 amphipathic helix residues (MID-F758, M762). Proteins colored as in Figure 1A. **(C**, **D)** Co-immunoprecipitation analyses in HEK293 cells showing the association between MHF and MID variants as indicated. Asterisks indicate cross-reactive non-specific polypeptides (IgG light chains). Because the MID domain is the only domain in FANCM that binds MHF [14], a mutation that disrupts MID-MHF association should have the same effect in the context of the full-length protein. **(E**, **F)** Immunoblotting shows the levels of FANCD2 and FANCM in whole-cell lysates from various DT40 cells as indicated, which include: wild-type (WT), *FANCM*$^{-/-}$ cells, and *FANCM*$^{-/-}$ cells complemented with human WT or mutants of FANCM. The ratio between the monoubiquitinated and non-ubiquitinated FANCD2 (L/S) is shown. Cells were treated with MMC (50 ng/ml) for 18 h. BAF57 or actin was used as a loading control. **(G**, **H)** Histograms showing spontaneous SCE levels of various DT40 cells as indicated. The mean number of SCEs per metaphase is listed. *P*-values were calculated using the Student's *t*-test.

(V749G/H751G and F758A/M762A) based on co-immunoprecipitation assays. All variants fractionated with profiles similar to that of the wild-type protein, with peak fractions near 34, which roughly corresponds to the calculated molecular weight of Flag-MID-MHF complex (about 70 KDa) (Supplementary information, Figure S5). No variants fractionated with peaks corresponding to the calculated molecular weight of Flag-MID alone (about 17 KDa), which is likely due to instability of MID without its partners (recombinant MID in *Escherichia coli* was also unstable without co-expression of MHF1 and MHF2; data not shown). These data suggest that all MID variants can fold normally and form complexes with MHF in the extracts. However, the conformation of MID-MHF complexes formed by V749G/H751 and F758A/M762A could be somewhat different from that of the wild type, because their peak fractions were slightly shifted (from 32-34 in wild type to 34 in mutants). In addition, V749G/H751G and its associated MHF1 also fractionated with an extra peak (fraction 28), which was not observed in the wild-type complex and could therefore represent an aberrant complex. Combined with co-immunoprecipitation results, these data suggest that V749G/H751 and F758A/M762A can form complexes with MHF in crude extracts, but these complexes are unstable and cannot withstand the washing conditions of co-immunoprecipitation assays.

In addition to studies of the central motifs, we also mutated several residues in the peripheral motifs of MID that wrap around MHF and engage in both hydrophobic and polar interactions (L680/W688 and L785/D789; Supplementary information, Figure S4B and S4C). None of these mutations significantly affected the association between Flag-tagged MID and MHF (Figure 2D), arguing that interactions through the peripheral motifs of MID are less important compared to those through the central motifs.

The Zn coordination structure is required by FANCM to promote normal activation of the FA network and suppress SCE

To study the functional relevance of the newly identified Zn coordination structure, we examined whether FANCM variants carrying mutations in Zn liganding residues, H751G and C755G, are defective in activation of the FA network and suppression of SCEs, using FANCM-knockout DT40 cells as described previously [7, 14]. Consistent with previous data, *FANCM*$^{-/-}$ DT40 cells displayed reduced levels of monoubiquitinated FANCD2 in response to mitomycin C (MMC) (Figure 2E and 2F, lanes 1-2), as well as higher levels of SCEs (Figure 2G and 2H). These two abnormal phenotypes were largely corrected by re-introduction of wild-type FANCM (Figure 2E and 2F, lanes 1-3; Figure 2G and 2H), but not by FANCM-H751G or -C755G mutants (Figure 2E and 2G). We also examined FANCM-V749G/H751G double mutant, which is defective in both Zn coordination and hydrophobic interactions; and found that it also failed to correct the two abnormal phenotypes (Figure 2F and 2H). These data indicate that the Zn structure is critical for FANCM to promote normal activation of the FA network and suppress SCE.

Identification of the residues in MHF that are essential for its interaction with FANCM

The most conserved residues in MHF are those tracing their interface with MID (Supplementary information, Figure S2A and S2B). We mutated several such residues to determine whether they affect interactions with MID (Supplementary information, Table S2). Because MHF1 and MHF2 are present in two copies in the MID-MHF complex, and each copy can interact with MID, mutation of one MHF residue could affect interactions at two symmetrical locations.

For MHF1, a pairwise point mutant, A48Q/E52R, was designed to disrupt the interactions between the MHF1 A-chain and MID α3-α4, and between the MHF1 C-chain and MID β1-L2 (Figure 3A and data not shown). This mutant protein co-immunoprecipitated with a drastically reduced level of FANCM (less than 20%), but with normal level of MHF2 when transfected into *MHF1*$^{-/-}$ DT40 cells (Figure 3B). The data indicate that these MHF1 residues are important for formation of a stable MID-MHF complex, but not the MHF heterotetramer. For MHF2, the single point mutation P75Q was generated to sterically hinder the MID α2 helix as well as the residues L733 and F758 from protruding into the hydrophobic cleft between the MHF2 α3 C-terminal helices (Figure 3C). This mutant co-immunoprecipitated with undetectable level of FANCM, but normal level of MHF1 (Figure 3D), suggesting that P75 is also critical for stable association between MHF and FANCM. Combined with MID mutational data, these results suggest that MID and MHF residues within the central "core" of the MID-MHF interface (between MID L3 loop and α3 helix (Figure 1A)) provide crucial interactions for stabilizing the MID-MHF complex.

We have also mutated MHF2-D80 residue, which was predicted to engage in Zn coordination with MID (Figure 1B), as well as an intramolecular interaction with the MHF2-R52 residue (data not shown). However, while D80 mutation did disrupt MHF2-FANCM association, it also destabilized the MHF1-MHF2 complex (data not shown). We were unable to distinguish whether its effect

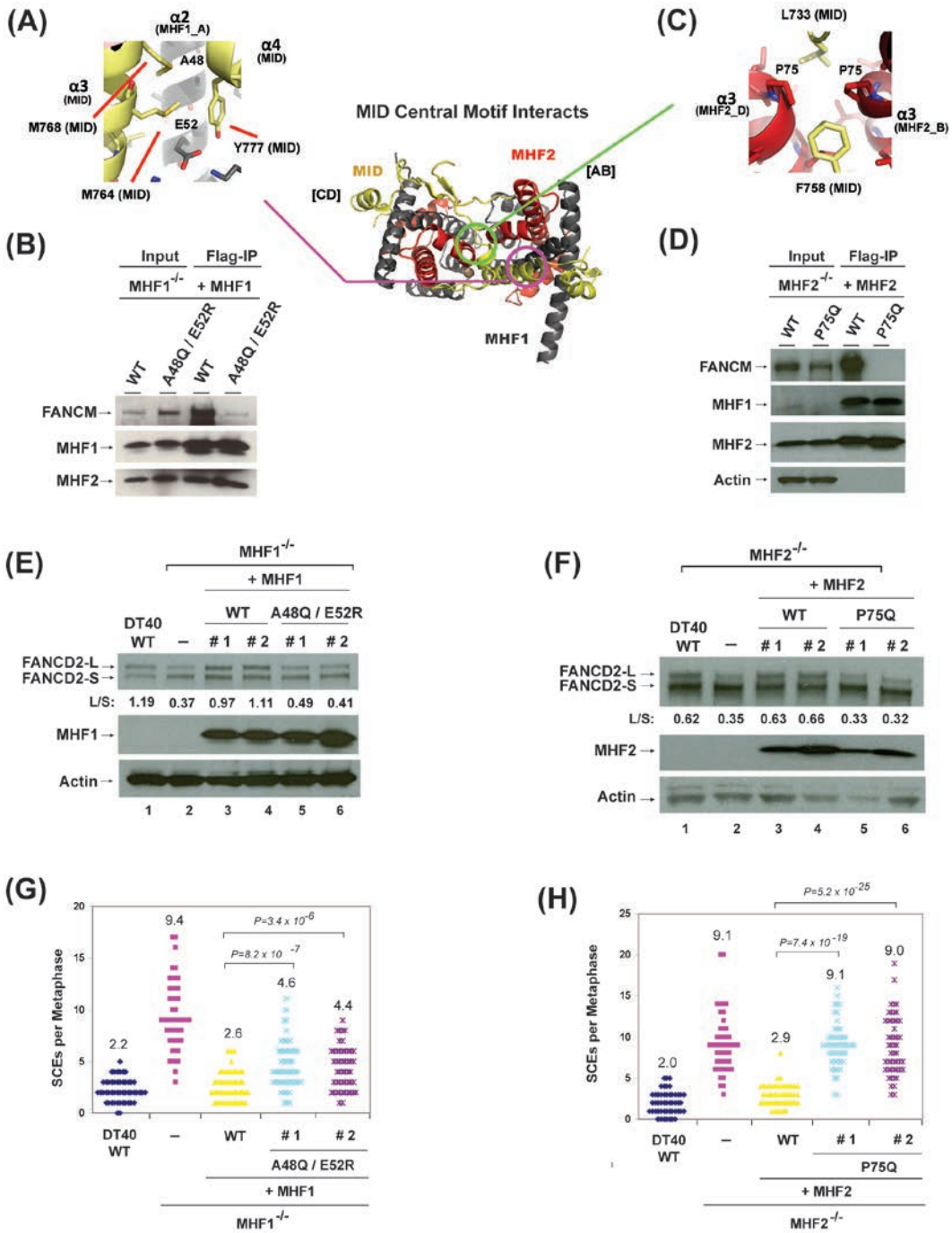

Figure 3 Identification of the residues in MHF that are essential for stabilization of MID-MHF complex, normal activation of the FA network and suppression of SCEs. **(A**, **C)** Key MID and MHF interfaces are shown in detail. MID central motif residues: **(A)** MID residues at the MHF1-A48/E52 interface (MID-M764, M768, Y777), **(C)** MID residues at the MHF2-α3 helix P75 interface (MID-L733, L758). Proteins colored as in Figure 1A. **(B**, **D)** Co-immunoprecipitation analyses showing association between full-length FANCM and different variants of MHF in chicken DT40 cells mutated of either MHF1 **(B)** or MHF2 **(D)**. **(E**, **F)** Immunoblotting shows the levels of FANCD2, MHF1 and MHF2 in whole-cell lysates from various DT40 cells as indicated, which include: wild-type (WT), *MHF1*$^{-/-}$ cells, *MHF2*$^{-/-}$ cells, and their complemented versions using human WT or mutants of MHF1 or MHF2. The ratio between the monoubiquitinated and non-ubiquitinated FANCD2 (L/S) is shown. Cells were treated with MMC (50 ng/ml) for 18 h. Actin was used as a loading control. **(G**, **H)** Histograms showing spontaneous SCE levels of various DT40 cells as indicated. The mean number of SCEs per metaphase is listed. *P*-values were calculated using the Student's *t*-test.

on MHF2-FANCM association is due to disruption of Zn coordination or destabilization of the MHF complex.

Stable interaction between FANCM and MHF are required for normal activation of the FA network and suppression of SCEs

To further investigate whether interactions between MHF and FANCM through the newly identified interface are important *in vivo*, we examined whether MHF1-A48Q/E52R and MHF2-P75Q mutants, which are defective in FANCM association, are compromised in activatation of the FA network and suppression of SCEs. Consistent with earlier findings [14], chicken DT40 cells inactivated of either MHF1 or MHF2 displayed reduced levels of monoubiquitinated FANCD2 in response to MMC (Figure 3E and 3F, lanes 1-2), as well as higher levels of SCEs (Figure 3G and 3H, lanes 1-2). Re-introduction of wild-type version of each protein largely corrected these two abnormal phenotypes (Figure 3E and 3F, lanes 1-4; 3G and 3H), whereas re-expression of MHF2-P75Q and MHF1-A48Q/E52R mutants was largely ineffective in correcting these phenotypes (Figure 3E-3H). For example, whereas wild-type MHF1 restored the ratio between monoubiquitinated and non-ubiquitinated FANCD2 from 0.37 to 1.04 (average of two independent clones, herein after the same), the MHF1-A48Q/E52R mutant restored this ratio to only 0.45, about 12% as active as the wild-type protein (Figure 3E). Similarly, whereas wild-type MHF1 suppressed SCE frequency of *MHF1*$^{-/-}$ cells from 9.4 to 2.6, the MHF1-A48Q/E52R mutant suppressed the frequency to 4.5, about 30% less active than the wild-type protein (Figure 3G). Collectively, these data demonstrate that the interactions between MHF and FANCM through the newly identified interfaces are required for the proper function of the FANCM-MHF complex in activation of the FA network and suppression of SCEs.

One likely explanation for the above data is that normal activation of the FA network and suppression of SCEs depend on the DNA binding and remodeling activities of FANCM [10, 18]; and these two activities are strongly stimulated by MHF [14, 19]. The various mutations above disrupt interactions between FANCM and MHF to various degrees, which likely compromise the DNA binding and remodeling activities of FANCM-MHF, leading to defective activation of the FA network and SCE suppression.

The FANCM-MHF complex is structurally and functionally independent of the CENP-T-W-S-X complex

A previous report showed that GFP-tagged FANCM localizes at centromeres, and this localization depends on MHF [22]. Because the centromere localization of MHF (CENP-S-X) depends on CENP-T-W through formation of the CENP-T-W-S-X complex [20, 25], these findings raised a possibility that FANCM may associate with the CENP-T-W-S-X complex and mediate the assembly of kinetochores. To examine this possibility, we fitted the MID of FANCM onto the CENP-T-W-S-X heterotetramer by structural modeling, and found that the resulting MID-CENP-T/W/S/X pentamer is likely to be unstable because it cannot coordinate the structural zinc atom that is necessary for stabilization of the MID-MHF pentamer (Supplementary information, Figure S6). Consistent with this prediction, no CENP proteins, except MHF1 and MHF2, were detected in the complex co-immunoprecipitated with FANCM (Figure 4A-4C); whereas four CENP proteins, including CENP-T, co-immunoprecipitated with MHF1 (Figure 4A-4C), indicating that FANCM associates with MHF but not CENP-T-W-S-X.

We found that the localization and signal intensities of an outer kinetochore protein, Ndc80, at kinetochores of *FANCM*$^{-/-}$ DT40 cells were indistinguishable compared to the wild-type cells (Figure 4D). This was in contrast to the earlier reports that Ndc80 exhibited reduced kinetochore localization in DT40 cells deficient in MHF1 (CENP-S), MHF2 (CENP-X) or CENP-T [20, 25]. In addition, whereas the localization of MHF1 (CENP-S) to kinetochores is reduced in *CENP-T*$^{-/-}$ and *CENP-W*$^{-/-}$ cells [23], the knockout of FANCM had no impact on the MHF1 kinetochore localization (Figure 4D). Furthermore, the centromere localization of CENP-T was also unaffected in *FANCM*$^{-/-}$ cells (Figure 4D). Taken together with the biochemistry evidence above, these data suggest that FANCM associates only with MHF, and the resulting FANCM-MHF complex does not participate in kinetochore assembly.

Because CENP-T, along with its partner CENP-W, associates with MHF1 (CENP-S) and MHF2 (CENP-X), we studied the possibility whether CENP-T may play a role in the FA network and suppression of SCEs. We found that the level of monoubiquitinated FANCD2 in *CENP-T*$^{-/-}$ DT40 cells was indistinguishable from that of wild-type cells with or without MMC treatment (Figure 4E), and the levels of SCEs were also comparable between wild-type and *CENP-T*$^{-/-}$ DT40 cells (Figure 4F). These data are in contrast to the results of DT40 cells inactivated of MHF1, MHF2 or FANCM (Figures 2E and 2F, 3E and 3F) [7, 14]. The data indicate that, of the two MHF-containing complexes, only FANCM-MHF participates in the FA network and suppression of SCEs, whereas the CENP-T-W-S-X complex does not (Figure 4G).

Asymmetric and allosteric remodeling of MHF tetramer by MID

The MHF tetramers with or without the bound MID

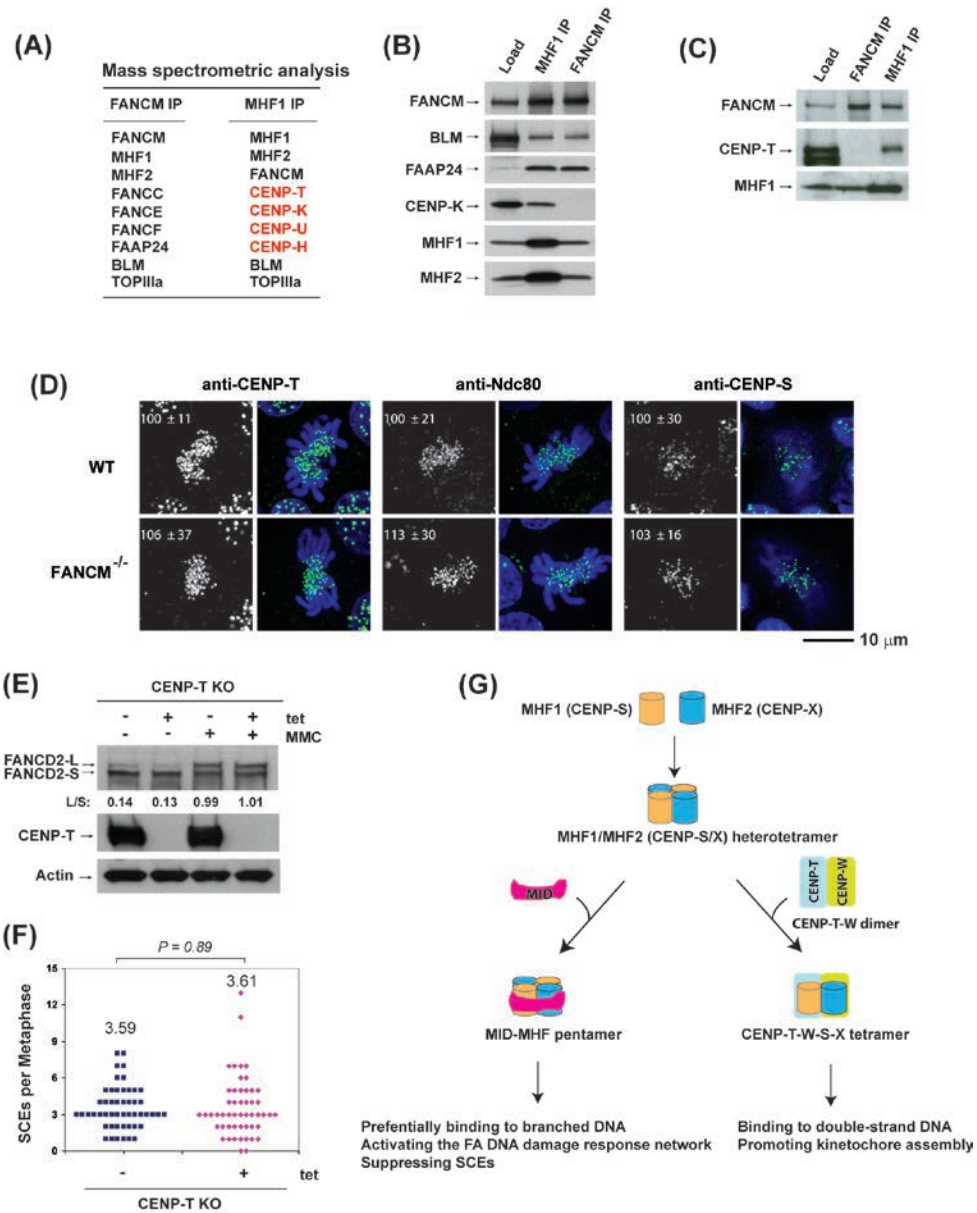

Figure 4 MHF1 and MHF2 are present in two functionally distinct complexes. **(A)** Mass spectrometric analysis shows that MHF1 (CENP-S) and MHF2 (CENP-X) are present in both FANCM and MHF1 immunoprecipitates isolated from the chromatin fraction of HeLa cells by either FANCM or MHF1 antibody, whereas other centromere proteins (CENP-T, CENP-K, CENP-U, CENP-H) were detected only in the MHF1 immunoprecipitate. **(B**, **C)** Immunoblotting shows that MHF1, MHF2, FANCM and other components of the FA core and Bloom complexes were present in both FANCM and MHF1 immunoprecipitates, but CENP-K and CENP-T were detected only in the MHF1 immunoprecipitate. **(D)** Immunofluorescence analyses of wild-type (WT) or *FANCM*$^{-/-}$ DT40 cells with different antibodies (green) as indicated. DNA was stained by DAPI (blue). The signal intensities of CENP-T, Ndc80 and CENP-S at each kinetochore detected by immunofluorescence were measured relative to an adjacent background signal ($n > 10$ cells; average intensity ± SD). The scale bar represents 10 μm. **(E)** Immunoblotting shows the levels of FANCD2 and CENP-T in CENP-T conditional knockout (KO) DT40 cells with or without the treatment of tetracycline (tet) for 24 h. The ratio between the monoubiquitinated and non-ubiquitinated FANCD2 (L/S) is shown. Cells were treated with MMC (50 ng/ml) for 16 h. Actin was used as a loading control. **(F)** Histograms showing the levels of spontaneous SCEs in CENP-T conditional KO DT40 cells with or without the treatment of tet for 24 h. The mean number of SCEs per metaphase is listed. *P*-value was calculated using the Student's *t*-test. **(G)** A diagram describes that MHF has two distinct functions by forming independent complexes with FANCM and CENP-T/W. One is the FANCM-MHF complex that preferentially binds branched DNA and functions specifically in DNA repair. The other is CENP-T-W-S-X complex that promotes kinetochore assembly at centromeres [20].

domain align with an overall root-mean-square deviation (RMSD) of 1.708 Å, suggesting that the MHF structure was remodeled upon MID binding.

We compared the distances between the identical atoms in the MHF tetramer vs MID-MHF pentamer by difference distance mapping (DDM) [26], and found that the distances changed substantially between the heterodimers but minimally within each heterodimer. Overall, the MHF [AB] heterodimer undergoes the least changes, 0.561 Å RMSD, and was used as the basis for structural comparisons in Figure 5.

The largest changes occur along the MHF [CD] surface directly contacted by MID: the MHF1 and MHF2 α2 helices move away by 6.9 Å and 5.5 Å, respectively, from the MID α2 helix, which traverses between the MHF heterodimers (Figure 5A). Lesser, yet still significant, changes occur along the MHF surface not directly contacted by MID, where the same two helices move by 1.4 Å and 4.2 Å, respectively, parallel to the other MHF heterodimer (Figure 5B). A kink develops upon MID binding at Cys58 of the MHF1 [CD] α2 helix (Figure 5C), but not at the symmetrical position in the MHF1 [AB] α2 helix (Figure 5D). Together, these data show that MID binding results in both allosteric and asymmetric remodeling of the MHF tetramer.

It should be mentioned that in the study by Tao *et al.* [22], MHF without and with MID aligned with an overall RMSD of 0.87 Å, which is smaller than 1.708 Å observed in the current study. We compared their data with ours and found that the structures of MID-MHF complexes are identical in both studies, whereas the structures of MHF tetramer are slightly different despite that each MHF1-MHF2 dimers are superimposable. The different MHF crystal structures may be due to the intrinsic flexibility of the heterotetramer.

MID alters the number of positive charges and allosteric structures in the MHF DNA-binding surfaces to favor branched DNA

Previous structural studies suggested that the DNA-binding surface of MHF is a positively charged electrostatic surface similar to that in histone H3-H4 tetramer, and mutations of these charged residues decrease the DNA-binding activity of MHF [20, 22]. Tao *et al.* [22] also suggested that the binding of MHF by MID creates a new DNA-binding surface consisting of positively charged residues from both proteins, and that MID increased the positive charges in one area of the surface, but decreased them in another area [22].

To identify residues in MHF or MID-MHF that bind DNA, we generated models for each dsDNA-bound complex by superimposing the MHF tetramer onto the DNA-bound nucleosome structure of the histone H3-H4 tetramer (PDB ID: 3AFA, Figure 6A). We also compared surface electrostatic maps generated for MHF and MID-MHF structures and identified additional potential DNA contacts (Figure 6A-6E; Supplementary information, Figures S2 and S3). Most of the residues identified match those previously suggested to contact DNA [20-22], including many positively charged residues on MHF1-MHF2 tetramer, at the C-terminal tail of MHF1, and a cluster from MID (Figure 6A). However, our model predicts that T76 of MHF1 and R695 of FANCM may also interact with DNA.

Notably, we found that MID changed the electrostatic surfaces along the predicted α1α1 DNA-binding surface of the MHF heterotetramer (Figure 6B). The α1α1 surface of the MHF [CD] heterodimer increases from five to seven positively charged residues (Figure 6C), while that of the MHF [AB] heterodimer decreases from five to zero positively charged residues (Figure 6D). These changes are expected to enhance MHF [CD] DNA binding at the α1α1 site, but disrupt MHF [AB] binding at the α1α1 site. In contrast, the MHF L1L2 DNA-binding surfaces, which contact two helical turns of DNA, remain largely unmodified by MID in terms of electrostatic charge; however, the allosteric structural changes of the MHF [CD] heterodimer caused by binding of MID (Figure 5B), alter the shape and trajectory of the DNA-binding surface (Figure 6E). Overall, the combined DNA-binding surface in the MID-MHF pentamer becomes asymmetric due to: (1) occlusion of the MHF [AB] α1α1 DNA-binding surface, (2) extension of the MHF [CD] α1α1 DNA-binding surface by MID and (3) distortion of the MHF [CD] heterodimer. We hypothesize that the shortened and asymmetric DNA-binding surface of the MID-MHF pentamer may disfavor regular duplex DNA, but rather, favor branch-point DNA. Consistent with this prediction, the MID-MHF complex had lower binding activity for dsDNA (Figure 6F, Supplementary information, Figure S7), but higher activity for a branched DNA — a model Holliday junction (HJ) (Figure 6G; Supplementary information, Figure S8A and S8B), compared to MHF alone that favors dsDNA as reported previously [14] (Figure 6F and 6G, left panels; Supplementary information, Figure S8C, S8D). These data support our hypothesis that MID remodels MHF structure to favor branched DNA.

Conserved lysine and arginine clusters from both MID and MHF contribute to binding of the MID-MHF complex to HJ DNA

The previous studies have used dsDNA-based binding assays to confirm the importance of the predicted

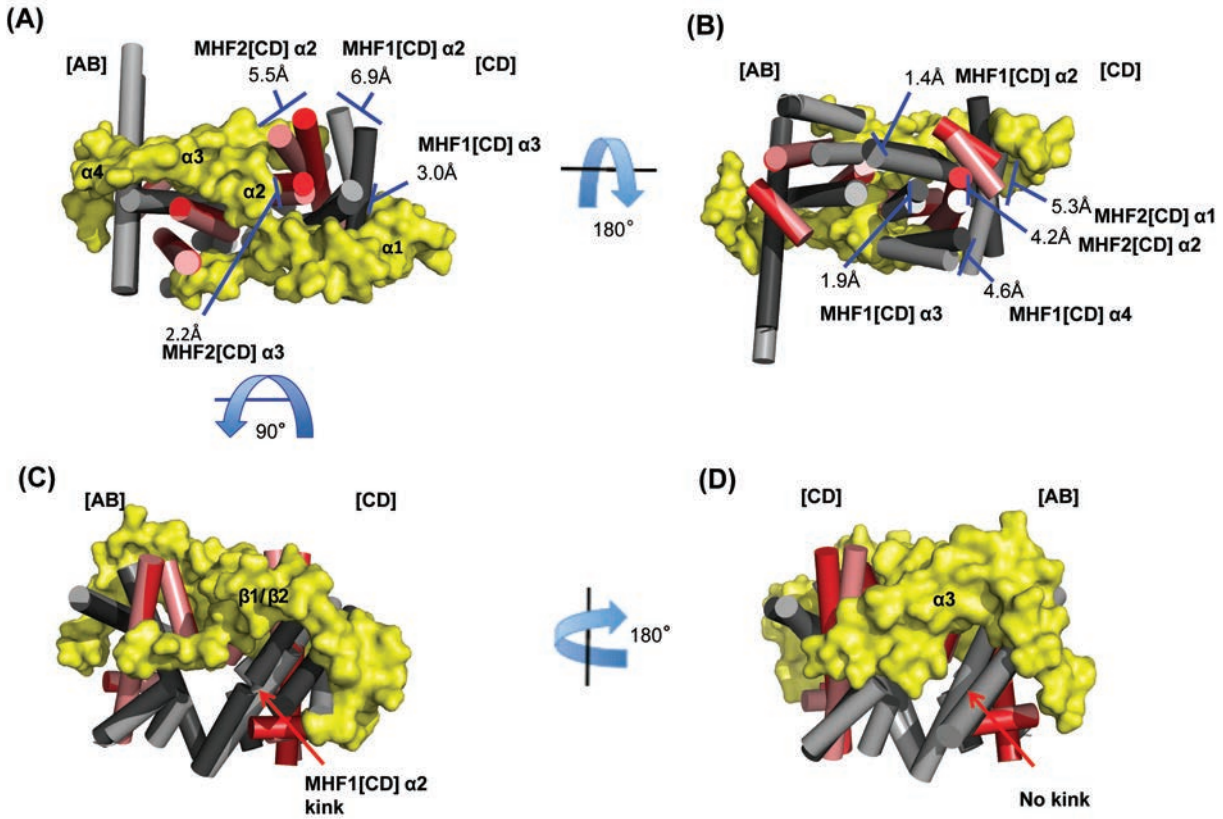

Figure 5 MID asymmetrically remodels the MHF tetramer. **(A-D)** A 360 degree view of the aligned structures. MHF (MHF1, light gray; MHF2, salmon) and MID-MHF (MHF1, dark gray; MHF2, red; MID, yellow) were aligned by the MHF [AB] heterodimer. Helices for the MHF heterotetramer have been shown as cylinders and MID is displayed as a surface (yellow). The largest differences in distance are observed in the [CD] MHF heterodimer, namely the MHF1 and MHF2 α2 helices. Distances between relative positions in the MHF and MHF-MID structures are labeled accordingly. The kink in MHF2 [CD] α2 helix has been labeled along with the secondary structure elements of MID.

DNA-interacting residues in MHF or MID-MHF [20, 22]. Because biochemistry and genetic data have suggested that branched DNA is the physiological substrate that FANCM-MHF binds and models [8, 9, 11, 14, 15], we used the HJ DNA as a substrate to test binding by recombinant MID-MHF complexes carrying substitutions of various positively charged residues at the predicted DNA-binding surfaces. Largely consistent with previous results from dsDNA-based assays [20, 22], all tested variants displayed either reduced binding activity or altered mobility in gel-shift assays (Figure 7A-7C, and Supplementary information, Figure S7). These data indicate that the identified surfaces from both MHF and MID are important for binding to HJ DNA *in vitro*.

The integrity of the DNA-binding surfaces of MID-MHF is required for normal activation of the FA network and suppression of SCEs

To study the importance of the newly identified DNA-binding surface of MID-MHF, we introduced several DNA-binding point mutants of each protein (Figure 7D-7G) into their respective knockout chicken DT40 cells. While the wild-type version of each protein fully restored the defective FANCD2 monoubiquitination and suppressed the higher level of SCEs, the MHF1-R70A mutant, which showed severe deficiency in DNA binding (Figure 7B), was completely defective in FANCD2 monoubiquitination (Figure 7D) and partially defective in suppression of SCES (Figure 7E). In addition, both MHF2-K27A/K29A and FANCM-R693A/R695A double mutants, which showed less severe DNA-binding deficiency (Figure 7A and 7C), were partially deficient in suppression of SCEs (Figure 7F and 7G). These data suggest the newly identified DNA interface is important for FANCM-MHF to promote activation of the FA network and protect genome integrity.

We noticed that the FANCM-R693A/R695A mutant suppressed the SCE frequency from 15.9 to 4.3 (Figure

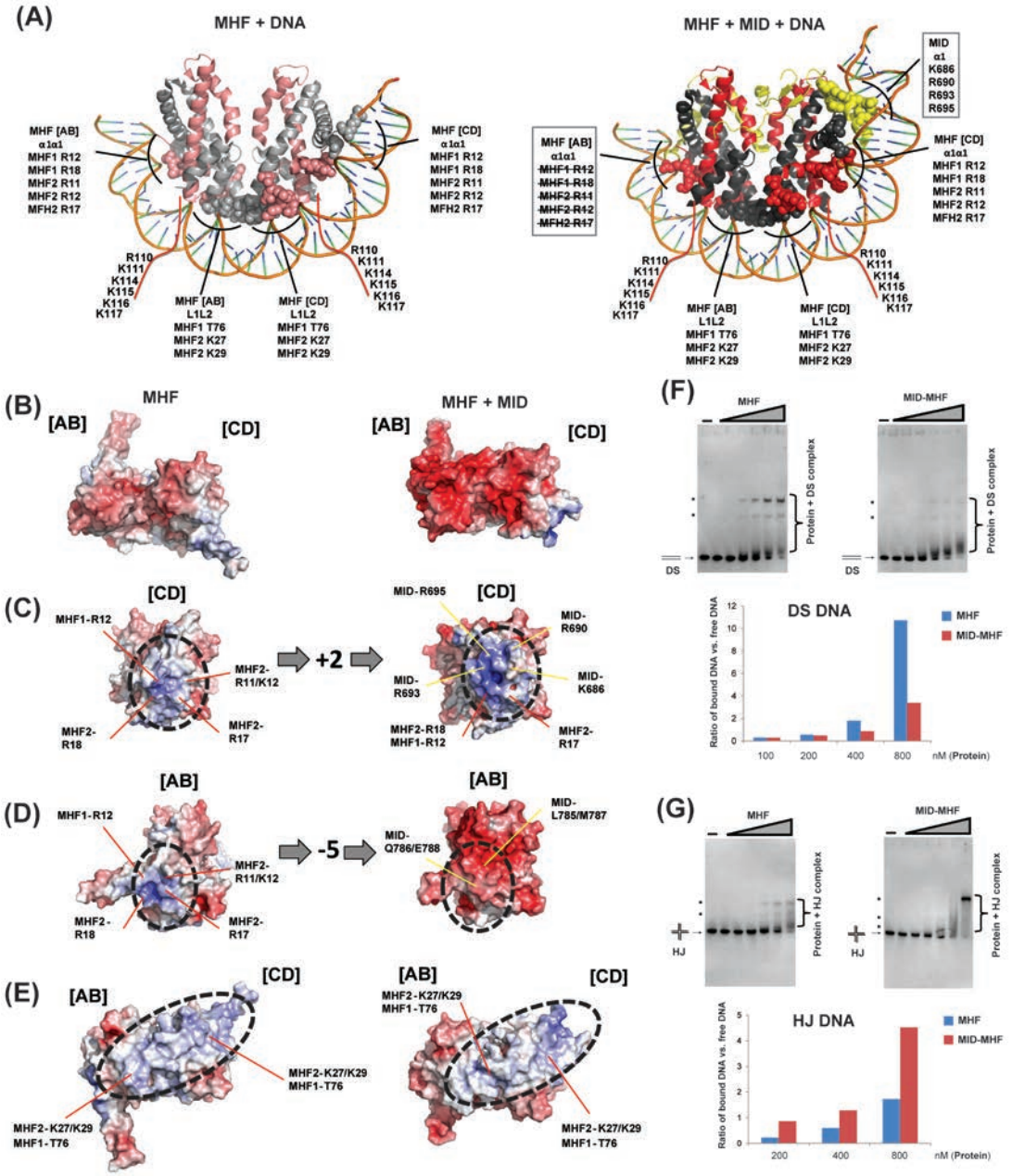

Figure 6 MID alters the number of positive charges and allosteric structures in the MHF DNA-binding surfaces to favor branched DNA. **(A)** Structural models of the interaction between MHF tetramer (left) or MID-MHF pentamer (right) and double-stranded DNA as seen in the histone H3-H4 nucleosome structure (PDB ID: 3AFA). Potential MHF and MID DNA-interacting residues are labeled for the $\alpha 1\alpha 1$ and L1L2 protein-DNA contact sites. The residues highlighted in the two boxes are altered upon MID association: those crossed by lines are occluded by MID, whereas the other ones are contributed by MID (see **B**-**D** below). Red lines indicate the unstructured MHF1 C-terminal tail with residues predicted to contact DNA labeled. **(B**-**E)** Left panels show the MHF heterotetramer. Right panels show the MID-MHF pentamer. Electrostatic surface potentials are shown (red; electronegative; white, neutral; blue, electropositive; +/–10 kcal/electron unit, PyMOL APBS [43]). The binding of MID increased the number of positive charges in the highlighted surface area of MHF [CD] **(C)**, decreased the positive charges in another area of MHF [AB] **(D)**, and caused allosteric changes in the areas marked by red lines **(E)**. **(F)** Top panel: EMSA for the comparison of the DNA-binding activity of MHF and MID-MHF complexes with double-stranded DNA (DS). The reaction contained the Digoxigenin-labeled DS DNA substrate (1 nM) with or without MHF and MID-MHF complexes as indicated. The protein concentrations are: 0, 25, 50, 100, 200, 400 and 800 nM, from left to right. Bottom: the specific protein-bound DNA bands (marked with dots) and free DS DNA substrate were quantified by KODAK Molecular Imaging Software, and their ratios are shown by a graph. **(G)** As described in **F**, excepted that HJ DNA was used.

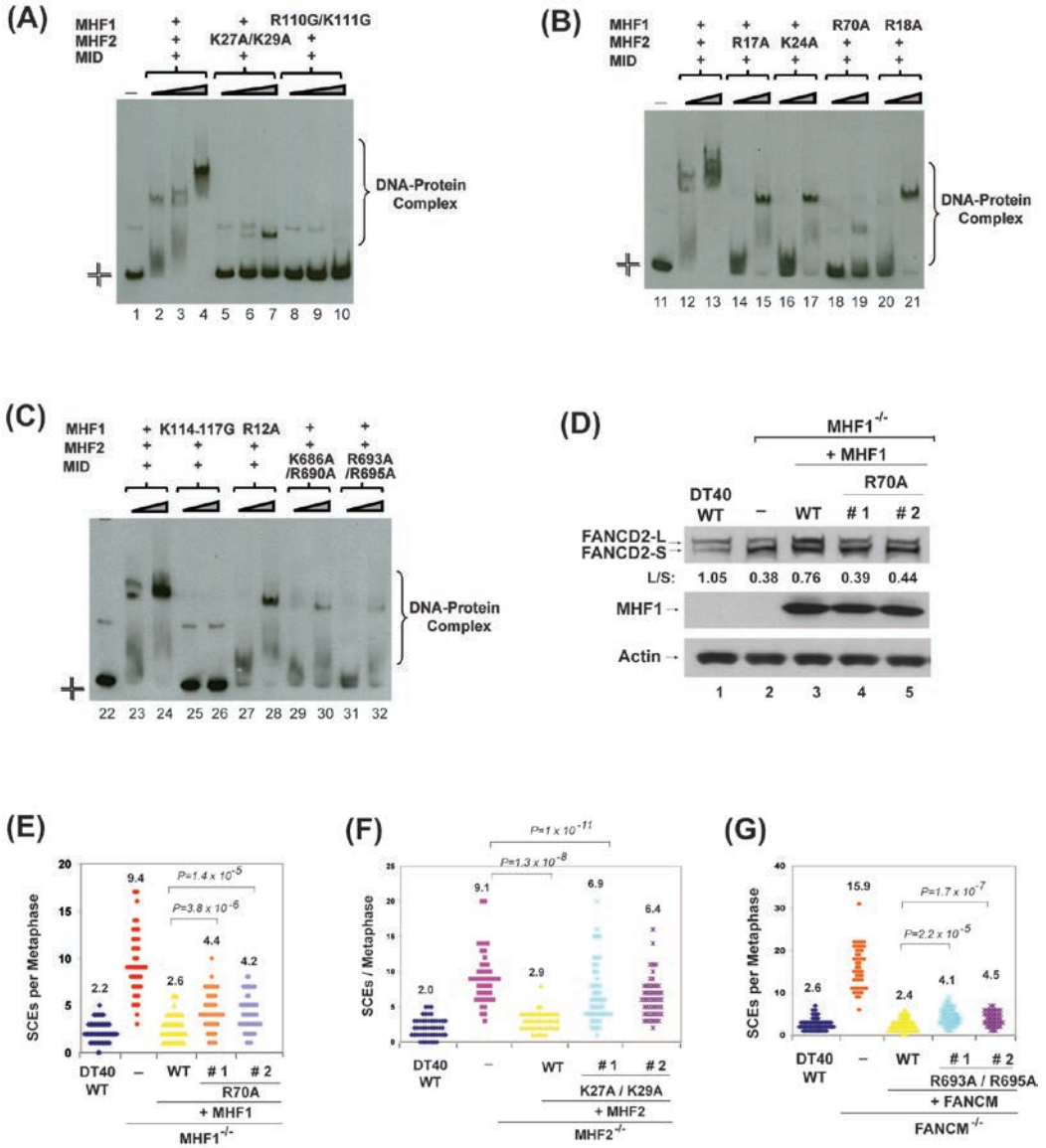

Figure 7 The integrity of the DNA-binding interfaces of MID-MHF complex is required for normal activation of the FA network and suppression of SCEs. **(A-C)** Comparison of the DNA-binding activity of MID-MHF complexes containing either wild type or mutants by EMSA. The reaction contained the Digoxigenin-labeled Holliday junction DNA substrate (1 nM) with or without wild type and mutants of the MID-MHF complex as indicated. The protein concentrations are: MID-MHF complex wild type: 200 nM in lane 2; 400 nM in lanes 3, 12 and 23; 800 nM in lanes 4, 13 and 24; MID-MHF complex carrying MHF2 point mutant (K27A/K29A): 200, 400 and 800 nM in lanes 5-7, respectively; MID-MHF complex carrying MHF1 point mutant (R110G/K111G): 200, 400 and 800 nM in lanes 8-10, respectively; MID-MHF complex carrying MHF2 point mutant (R17A): 400 and 800 nM in lanes 14 and 15, respectively; MID-MHF complex carrying MHF2 point mutant (K24A): 400 and 800 nM in lanes 16 and 17, respectively; MID-MHF complex carrying MHF1 point mutant (R70A): 400 and 800 nM in lanes 18 and 19, respectively; MID-MHF complex carrying MHF1 point mutant (R18A): 400 and 800 nM in lanes 20 and 21, respectively; MID-MHF complex carrying MHF1 point mutant (K114-117G): 400 and 800 nM in lanes 25 and 26, respectively; MID-MHF complex carrying MHF1 point mutant (R12A): 400 and 800 nM in lanes 27 and 28, respectively; MID-MHF complex carrying MID point mutant (K686A/R690A): 400 and 800 nM in lanes 29 and 30, respectively; MID-MHF complex carrying MID point mutant (R693A/R695A): 400 and 800 nM in lanes 31 and 32, respectively. **(D)** Immunoblotting shows levels of FANCD2 and MHF1 in whole-cell lysates from the indicated DT40 cells: wild-type (WT), *MHF1*$^{-/-}$ cells, and *MHF1*$^{-/-}$ cells complemented with human WT or mutant of MHF1. The ratio between the monoubiquitinated and unubiquitinated FANCD2 (L/S) was shown. Cells were treated with MMC (50 ng/ml) for 18 h. Actin was used as a loading control. **(E-G),** Histograms showing spontaneous SCE levels of various DT40 cells as indicated. The mean number of SCEs per metaphase is listed. *P*-values were calculated using the Student's *t*-test.

7G), which is about 85% as active as the wild-type protein (that suppressed SCE frequency from 15.9 to 2.4). In contrast, the MHF1-R70A and MHF2-K27A/R29A mutant are more defective — about 75% and 40% as active as their respective wild-type proteins in SCE suppression (the two mutants suppressed SCE from 9.4 to 4.3 or from 9.1 to 6.7; whereas the wild-type proteins suppressed SCE from 9.4 to 2.6 and from 9.1 to 2.9, respectively) (Figure 7E and 7F). The data imply that the positively charged residues at the composite DNA-binding surface of MID-MHF contribute unevenly with regard to function of FANCM-MHF: the residues contributed by MID may be less important than those by MHF1 and MHF2. These results are consistent with our earlier findings that the DNA-binding activity of MHF is required for normal activation of the FA network and SCE suppression [14].

Discussion

MHF is remodeled by FANCM to favor branched DNA over dsDNA

The FANCM-MHF complex remodels branched DNA, but how the two proteins work together to recognize their substrate is unclear. An earlier model proposed that the two proteins cooperatively bind different parts of a branched DNA: FANCM at the branch point and MHF at the surrounding duplex regions [14] (Supplementary information, Figure S9A). Here, we show that MID asymmetrically remodels the DNA-binding surface of MHF in both electrostatic surface potential and allosteric structure. Specifically, MID increased the net surface positive charges along the MHF [CD] α1α1 DNA-binding surface, occluded all surface positive charges along the MHF [AB] α1α1 DNA-binding surface, and distorted the structure of the surface. These alterations are expected to strengthen dsDNA binding in one area, but disrupt it in another, leading to an overall negative effect on dsDNA binding. Consistent with this prediction, the MID-MHF complex has lower binding activity for dsDNA than MHF alone, but higher binding activity for branched DNA over dsDNA. This is in contrast to MHF alone, which shows the opposite preference. These data thus favor an alternative model for recognition of branched DNA by FANCM-MHF (Supplementary information, Figure S9B), both the FANCM-helicase domain and the remodeled MID-MHF complex may bind the DNA branch point; and this binding is likely to be cooperative, as MHF can strongly enhance binding and remodeling activity of FANCM for branched DNA [14]. This model is supported by findings that point mutations at the composite DNA-binding surface from either protein weakened the binding activity of MID-MHF for a branched DNA substrate. It is also in agreement with the evidence that these mutations impaired the ability of FANCM-MHF to promote FANCD2 monoubiquitination, or to suppress SCEs, or both.

Remodeling of MHF by FANCM excludes it from association with the centromere complex

MHF is shared by two complexes with distinct functions: FANCM-MHF in DNA repair and CENP-T-S-W-X at centromeres. A study in yeast suggested that the two functions of MHF are distinct, and only the DNA repair function depends on FANCM [27]. However, a study in human cells revealed that GFP-tagged FANCM localizes at centromeres, and this localization depends on MHF [22]. Because the centromere localization of MHF (CENP-S-X) depends on CENP-T-W [25], these data imply that FANCM may associate with CENP-T/W/S/X to mediate kinetochore assembly in human cells. Nevertheless, our structural modeling and experiments disfavor this hypothesis. First, FANCM has no detectable association with any CENP proteins other than MHF1 and MHF2. Second, FANCM-deficient cells have no detectable defects in kinetochore assembly. Third, CENP-T-deficient cells exhibit normal activation of FA network and SCE suppression. Fourth, CENP-T/W/S/X lacks residues necessary to coordinate the intermolecular zinc atom with FANCM, which is important for stabilizing the FANCM-MHF complex. Our data support a scenario in which FANCM and CENP-T/W compete for a common pool of MHF to form two mutually exclusive complexes: the FANCM-MHF pentamer that excludes association with CENP-T/W and the CENP-T/W/S/X heterotetramer that cannot be stably bound by FANCM. This scenario may resemble the histone-fold-containing TAF complex that is also shared by multiple complexes (TFIID, SAGA and PCAF) [28]. We speculate that evolution may favor this scenario because existence of two separate complexes allows each one to evolve more efficiently to perform its specialized functions, which may be more advantageous than the presence of a single complex (FANCM-CENP-T/W/S/X) that does both jobs.

Remodeling of histone-fold complexes may create diversity in their DNA-binding affinity and specificity

Histone-fold complexes typically bind dsDNA and are often components of a variety of multiprotein complexes involved in DNA metabolism (nucleosomes, TFIID and FANCM-MHF), histone modification (SAGA, PCAF) [28] and kinetochore assembly (CENP-T-W-S-X). It is known that nucleosome structures are subject to covalent and non-covalent modifications by a variety of remodeling complexes. These modifications can weaken DNA-

histone interactions and produce altered or disassembled nucleosomes [29], which are subsequently recognized by other factors in transcription, replication and repair. In another case, a hsitone-fold-containing TAF complex in TFIID is altered by association with the TATA box-binding protein, TBP; and the resulting TFIID complex binds and bends dsDNA to allow assembly of transcription initiation complex [30]. Our findings that MHF is remodeled by FANCM represent a third scenario, in which a hsitone-fold complex is altered by its partner to favor branched DNA rather than dsDNA. We propose that MHF and other hsitone-fold complexes, such as nucleosomes, may have dynamic structures that can be readily adapted by their partners to create composite surfaces with altered DNA affinity or specificity to interact with biologically unique DNA structures generated during DNA metabolism.

While our manuscript was under revision, Zhao *et al*. [31] reported crystal structures of MHF-DNA complexes, and proposed a model that MHF prefers branched DNA over dsDNA. In contrast, we find that MHF prefers normal dsDNA both in this study and in an earlier publication [14]. This discrepancy might be due to different experimental conditions such as salt concentrations. Zhao *et al*. [31] used 150 mM KCl, whereas we used either phosphate saline [14] or 60 mM KCl (this study). Consistent with this notion, a previous study by several Zhao's colleagues [19] showed that the DNA preference of MHF appears to be affected by salt and protein concentrations. HJ DNA is preferred at 150 mM KCl; but at 100 mM KCl, the preference appears to change from HJ DNA to dsDNA when MHF protein concentration was increased [19]. Regardless of the discrepancy on MHF alone, Zhao *et al*. [31] showed that MID strongly stimulates binding of MHF to branched DNA, which is consistent with our finding that MID can remodel MHF structure to favor branched DNA.

Materials and Methods

Construct design, expression and purification

A bicistronic vector for co-expression of GST-tagged full-length human MHF1 (1-138 aa) and 6× HIS-tagged full-length human MHF2 (1-81 aa) in *E. coli* was generated by cloning MHF1 cDNA into the *Bam*HI/*Xho*I sites and MHF2 cDNA into *Nde*I/*Not*I sites of pGEX-BICIS-HIS [32]. For crystallization, MHF1 was truncated at residue R110. Human FANCM-MID (653-800) was cloned into pMBP-Parallel-1 based on pMal-C2 donating an N-terminal 6× HIS-MBP fusion tag. This construct was subsequently N-terminally truncated to FANCM-MID 669-800 for crystallization. Mutations and truncations were introduced into the above constructs by site-directed mutagenesis. MHF or MID-MHF constructs were transformed into BL21 (DE3) *E. coli* cells and grown at 30 °C until the OD_{600} reading reached 0.6. Protein expression was induced with 1 mM IPTG and cells were grown for 3 h prior to harvesting. MHF and MID-MHF-expressing cells were lysed in 50 mM sodium phosphate buffer, pH 8.0, containing 300 mM NaCl, 10% glycerol and lysozyme, protease inhibitors, $MgCl_2$, DNase and RNase. MHF alone was purified by Ni-Agarose and GST-Sepharose IMAC, and eluted with gradients of imidazole and glutathione sepharose, respectively. Both the 6× HIS- and GST-fusion tags were removed by incubation with thrombin overnight at 30 °C. MID-MHF was purified first by amylose resin and eluted with a maltose gradient, followed by MHF pruification as described above. The MBP-fusion tag was removed by incubation with TEV overnight at 4 °C. All samples were purified by size exclusion chromatography (Superdex 200 HR). Crystallization samples were prepared in 25 mM Tris, pH 7.5, 100 mM NaCl at 20-40 mg/ml.

Circular dichroism

Circular dichroism spectra were collected on a JASCO J-415 Spectrometer at the Biochemistry Core Facility, National Heart Lung and Blood Institute, National Institutes of Health. Spectra were collected at 222 nm between 5-90 °C every 0.5 °C. % Denatured was calculated as [(θ(5 °C)-θ(X °C))/(θ(5 °C)-θ(90 °C))].

Crystallization, data collection and refinement of MHF and MID-MHF complexes

Crystals of MHF (MHF1 1-110 and MHF2) and MID-MHF complexes were grown by sitting drop vapor diffusion at 295 K. Crystals, long- and thin-rod clusters, were initially grown in the sparse screen Wizard I (Emerald Bio) condition #28 containing 0.1 M Hepes, pH 7.5, 0.2 M NaCl, and 20% w/v PEG 3000. These crystals were optimized through generation of a seed stock from the original crystals that was used to seed into additional sparse matrix screens. The final conditions included 0.1 M Tris-HCl, pH 7.8, 0.2 M LiCl, 0.1 M Na_2SO_4 and 17.5% w/v PEG 3350. Crystals were cryoprotected in 10% glycerol and flash-frozen in liquid nitrogen. All datasets were collected at the advanced proton source (APS) beamline 23-ID-B equipped with a MAR 300 CCD detector. Diffraction data were reduced and scaled with XDS/XSCALE [33-35]. The MHF structure was solved by molecular replacement using PDB ID: 1TAF and the MID-MHF structure was solved with the previously solved MHF structure. All structures were refined using iterative cycles of TLS and restrained refinement with REFMAC5 [36], part of the CCP4 program suite [37], and model building was performed using the Crystallographic Object-Oriented Toolkit (Coot) [38]. All structures were validated using Molprobity [39] prior to deposition in the Protein Data Bank [40, 41]. Diffraction data and refinement statistics are listed for each structure (Supplementary information, Table S1).

Zinc analysis by the atomic absorption spectrometry

MHF and MID-MHF complexes were dialyzed against 12.5 mM Tris-HCl, pH 7.5, 50 mM NaCl. The dialyzed buffer was used to determine the background of ''Zn''. All proteins were diluted with 1% nitric acid as ''1/50'' and ''1/10''. Diluted protein solutions were boiled for 10 min to extract ''Zn'' and centrifuged to remove any precipitants. The supernatant of each sample was transferred to a fresh tube and ''Zn'' concentration was measured by a Perkin Elmer model 4100ZL atomic absorption spectrometer.

All measurements were done as triplicate for each sample and results were averaged of triplicate analyses.

Cell lines and antibodies

HEK293 cells, HeLa cells, *MHF*$^{-/-}$ DT40 cells, *FANCM*$^{-/-}$ DT40 cells and the antibodies against MHF1, MHF2, FANCM, FANCD2, FANCA, FAAP24, BLM, Actin and BAF57 have been previously described [14]. *MHF2(CENP-X)*$^{-/-}$ DT40 cells has been previously described [25]. CENP-T conditional knockout DT40 cells and anti-chicken CENP-S [25], CENP-T [23], and Ndc80 [42] antibodies have been previously described. Anti-Flag and anti-Flag M2 agarose beads were purchased from Sigma. Anti-CENP-K antibody was purchased from Abcam. Anti-human CENP-T antibody was purchased from Bethyl Laboratories, Inc.

Cell extraction, fractionation, gel filtration, immunoprecipitation and protein identification

Preparation of whole-cell extract, cell fractionation, Superose 6 gel filtration and immunoprecipitation were performed as described [14], except using both micrococcal nuclease and DNase I (2 U/μl) to treat the nuclear pellets. The eluted immunoprecipitates were subjected to mass spectrometry analyses for protein identification.

Plasmid construction, transfection and immunoblotting

pcDNA-Flag-MID was constructed by cloning a cDNA fragment of human FANCM (661-800 aa) with a Flag tag at its N-terminus into the *Nhe*I and *Bam*HI sites of pcDNA3.1. Various pcDNA-Flag-MID point mutants were generated by using the QuikChange kit (Stratagene). These constructed plasmids were transiently expressed in HEK293 cells, respectively. Cell extraction and immunoprecipitation with anti-Flag M2 agarose beads were performed as described [14]. The presence of various versions of MID, MHF1 and MHF2 was detected by immunoblotting.

Complementation, SCE and FANCD2 analyses

To perform complementation analysis, pcDNA3.1-Zeocin expression plasmid carrying wild type or mutants of chicken FANCM, Flag-tagged human MHF1 or Flag-tagged human MHF2 was transfected into *FANCM*$^{-/-}$ DT40 cells, *MHF1*$^{-/-}$ DT40 cells or *MHF2*$^{-/-}$ DT40 cells. In total, 0.5 mg/ml Zeocin was used for selection. Western blotting was performed to select the clones that stably express protein. The measurement of SCE and FANCD2 levels in various DT40 cells were performed as described [14].

Electrophoretic mobility shift assay (EMSA)

The indicated amounts of proteins were incubated with 1 nM of digoxigenin-labeled DNA substrates in 25 mM Tris-HCl, pH 7.5, 60 mM KCl, 5% glycerol, 2 mM $MgCl_2$ for 15 min at room temperature. The reactions were loaded on 4%-12% TBE gel (Invitrogen) and run in 0.5× TBE at 4 °C. Labeled DNA products were visualized by using DIG Wash and Block Buffer Set (Roche). The dsDNA substrate was generated by annealing oligo 1 and its complement (oligo 5). The HJ DNA substrate was generated by combining oligo 1, oligo 2, oligo 3 and oligo 4, as described [32]. Oligo 1 was labeled with digoxigenin at its 3′-end. The sequence of each oligo is listed below: Oligo 1: 5′-GTGACCGTCTCCGGGAGCTGGAAACGCGCGAGACGAAAGG-3′; Oligo 2: 5′-CCTTTCGTCTCGCGCGTTTCGCCAGCCCCGACACCCGCCA-3′; Oligo 3: 5′-TGGCGGGTGTCGGGGCTGGCCGCCGGGCAAGAGCAACTCG-3′; Oligo 4: 5′-CGAGTTGCTCTTGCCCGGCGCAGCTCCCGGAGACGGTCAC-3′; Oligo 5: 5′-CCTTTCGTCTCGCGCGTTTCCAGCTCCCGGAGACGGTCAC-3′.

Immunofluorescence analysis

DT40 cells were collected onto slides with a cytocentrifuge and fixed in 3% paraformaldehyde or cold methanol in phosphate-buffered saline (PBS) for 15 min and permeabilized in 0.5% NP-40 in PBS for 10 min at room temperature or 100% methanol for 15 min at −20 °C, rinsed three times in 0.5% BSA in PBS, and incubated for 1 h at 37 °C with primary antibody. Binding of primary antibody was then detected with FITC-conjugated goat anti-rabbit IgG (Jackson ImmunoResearch, West Grove, PA, USA) diluted to an appropriate concentration in 0.5% BSA in PBS. Affinity-purified rabbit polyclonal antibodies were used against recombinant chicken CENP-S, CENP-T and Ndc80. All subsequent analysis and processing of images were performed using Metamorph software (Molecular Devices, Japan). Kinetochore fluorescence intensities were determined by measuring the integrated fluorescence intensity within a 4 × 4 pixel square positioned over a single kinetochore and subtracting the background intensity of a 4 × 4 pixel square positioned in a region of cytoplasm lacking kinetochores. Maximal projected images were used for these measurements.

Accession numbers

The crystal structural data for MHF and MID-MHF have been deposited at the Protein Data Bank (PDB ID: 4E44 and 4E45).

Acknowledgments

This work was supported in part by the Intramural Research Program of the National Institute on Aging (AG000688-07), National Institute of Health; and by a Grant-in-Aid for Scientific Research (S) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan to TF.

References

1 Wang W. Emergence of a DNA-damage response network consisting of Fanconi anemia and BRCA proteins. *Nat Rev Genet* 2007; **8**:doi:10.1038/nrg2159.

2 Kim H, D'Andrea AD. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes Dev* 2012; **26**:1393-1408.

3 Kottemann MC, Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature* 2013; **493**:356-363.

4 Bogliolo M, Schuster B, Stoepker C, *et al*. Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *Am J Hum Genet* 2013; **92**:800-806.

5 Kashiyama K, Nakazawa Y, Pilz DT, *et al.* Malfunction of nuclease ERCC1-XPF results in diverse clinical manifestations and causes cockayne syndrome, xeroderma pigmentosum, and Fanconi anemia. *Am J Hum Genet* 2013; **92**:807-819.

6 Meetei AR, Medhurst AL, Ling C, *et al.* A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet* 2005; **37**:958-963.

7 Mosedale G, Niedzwiedz W, Alpi A, *et al.* The vertebrate Hef ortholog is a component of the Fanconi anemia tumor-suppressor pathway. *Nat Struct Mol Biol* 2005; **12**:763-771.
8 Gari K, Decaillet C, Delannoy M, Wu L, Constantinou A. Remodeling of DNA replication structures by the branch point translocase FANCM. *Proc Natl Acad Sci USA* 2008; **105**:16107-16112.
9 Gari K, Decaillet C, Stasiak AZ, Stasiak A, Constantinou A. The Fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks. *Mol Cell* 2008; **29**:141-148.
10 Xue Y, Li Y, Guo R, Ling C, Wang W. FANCM of the Fanconi anemia core complex is required for both monoubiquitination and DNA repair. *Hum Mol Genet* 2008; **17**:1641-1652.
11 Sun W, Nandi S, Osman F, *et al.* The FANCM ortholog Fml1 promotes recombination at stalled replication forks and limits crossing over during DNA double-strand break repair. *Mol Cell* 2008; **32**:118-128.
12 Blackford AN, Schwab RA, Nieminuszczy J, Deans AJ, West SC, Niedzwiedz W. The DNA translocase activity of FANCM protects stalled replication forks. *Hum Mol Genet* 2012; **21**:2005-2016.
13 Luke-Glaser S, Luke B, Grossi S, Constantinou A. FANCM regulates DNA chain elongation and is stabilized by S-phase checkpoint signalling. *EMBO J* 2010; **29**:795-805.
14 Yan Z, Delannoy M, Ling C, *et al.* A histone-fold complex and FANCM form a conserved DNA-remodeling complex to maintain genome stability. *Mol Cell* 2010; **37**:865-878.
15 Daee DL, Ferrari E, Longerich S, *et al.* Rad5-dependent DNA repair functions of the *Saccharomyces cerevisiae* FANCM protein homolog Mph1. *J Biol Chem* 2012; **287**:26563-26575.
16 Wang Y, Leung JW, Jiang Y, *et al.* FANCM and FAAP24 maintain genome stability via cooperative as well as unique functions. *Mol Cell* 2013; **49**:997-1009.
17 Singh TR, Bakker ST, Agarwal S, *et al.* Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. *Blood* 2009; **114**:174-180.
18 Rosado IV, Niedzwiedz W, Alpi AF, Patel KJ. The Walker B motif in avian FANCM is required to limit sister chromatid exchanges but is dispensable for DNA crosslink repair. *Nucleic Acids Res* 2009; **37**:4360-4370.
19 Singh TR, Saro D, Ali AM, *et al.* MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM. *Mol Cell* 2010; **37**:879-886.
20 Nishino T, Takeuchi K, Gascoigne KE, *et al.* CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. *Cell* 2012; **148**:487-501.
21 Yang H, Zhang T, Tao Y, *et al. Saccharomyces cerevisiae* MHF complex structurally resembles the histones (H3-H4) (2) heterotetramer and functions as a heterotetramer. *Structure* 2012; **20**:364-370.
22 Tao Y, Jin C, Li X, *et al.* The structure of the FANCM-MHF complex reveals physical features for functional assembly. *Nat Commun* 2012; **3**:782.
23 Hori T, Amano M, Suzuki A, *et al.* CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell* 2008; **135**:1039-1052.
24 ACOG Committee Opinion No. 469: Carrier screening for fragile X syndrome. *Obstet Gynecol*; **116**:1008-1010.
25 Amano M, Suzuki A, Hori T, *et al.* The CENP-S complex is essential for the stable assembly of outer kinetochore structure. *J Cell Biol* 2009; **186**:173-182.
26 Richards FM, Kundrot CE. Identification of structural motifs from protein coordinate data: secondary structure and first-level supersecondary structure. *Proteins* 1988; **3**:71-84.
27 Bhattacharjee S, Osman F, Feeney L, Lorenz A, Bryer C, Whitby MC. MHF1-2/CENP-S-X performs distinct roles in centromere metabolism and genetic recombination. *Open Biol* 2013; **3**:130102.
28 Albright SR, Tjian R. TAFs revisited: more data reveal new twists and confirm old ideas. *Gene* 2000; **242**:1-13.
29 Liu N, Balliano A, Hayes JJ. Mechanism(s) of SWI/SNF-induced nucleosome mobilization. *Chembiochem* 2011; **12**:196-204.
30 Elmlund H, Baraznenok V, Linder T, *et al.* Cryo-EM reveals promoter DNA binding and conformational flexibility of the general transcription factor TFIID. *Structure* 2009; **17**:1442-1452.
31 Zhao Q, Saro D, Sachpatzidis A, *et al.* The MHF complex senses branched DNA by binding a pair of crossover DNA duplexes. *Nat Commun* 2014; **5**:2987.
32 Ciccia A, Ling C, Coulthard R, *et al.* Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM. *Mol Cell* 2007; **25**:331-343.
33 Kabsch W. Automatic indexing of rotation diffraction patterns. *J Appl Cryst* 1988; **21**:67-72.
34 Kabsch W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constrants. *J Appl Cryst* 1993; **26**:795-800.
35 Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr* 2010; **66(Pt2)**:125-132.
36 Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 1997; **53**:240-255.
37 Winn MD, Ballard CC, Cowtan KD, *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 2011; **67**:235-242.
38 Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004; **60**:2126-2132.
39 Chen VB, Arendall WB III, Headd JJ, *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 2010; **66**:12-21.
40 Berman H, Henrick K, Nakamura H. Announcing the worldwide Protein Data Bank. *Nat Struct Biol* 2003; **10**:980.
41 Berman HM, Westbrook J, Feng Z, *et al.* The Protein Data Bank. *Nucleic Acids Res* 2000; **28**:235-242.
42 Hori T, Haraguchi T, Hiraoka Y, Kimura H, Fukagawa T. Dynamic behavior of Nuf2-Hec1 complex that localizes to the centrosome and centromere and is essential for mitotic progression in vertebrate cells. *J Cell Sci* 2003; **116**:3347-3362.
43 Schrodinger L. The PyMOL Molecular Graphics System, Version 0.99.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)