

Published in final edited form as:

Nat Commun. 2014 ; 5: 2987. doi:10.1038/ncomms3987.

MHF complex senses branched DNA via binding a pair of crossover DNA duplexes

Qi Zhao^{1,*}, Dorina Saro^{1,*}, Aristidis Sachpatzidis¹, Thiyam Ramsing Singh⁵, Daniel Schlingman¹, Xiao-Feng Zheng¹, Andrew Mack², Miaw-Sheue Tsai⁴, Simon Mochrie^{2,3}, Lynne Regan¹, Amom Ruhikanta Meetei⁵, Patrick Sung^{1,†}, and Yong Xiong^{1,†}

¹Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06520, USA

²Department of Applied Physics, Yale University, New Haven, Connecticut 06511 USA

³Department of Physics, Yale University, New Haven, Connecticut 06511 USA

⁴Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

⁵Division of Experimental Hematology and Cancer Biology, Cincinnati Children's Research Foundation and University of Cincinnati College of Medicine, Cincinnati, OH 45229, USA

Abstract

The conserved MHF1-MHF2 (MHF) complex functions in the activation of the Fanconi anemia (FA) pathway of DNA damage response, in regulating homologous recombination, and in DNA replication fork maintenance. MHF facilitates the processing of multiple types of branched DNAs by the FA DNA translocase FANCM. Here we report the crystal structure of a human MHF-DNA complex that reveals the DNA binding mode of MHF. The structure suggests an MHF preference for branched DNA over double stranded DNA through engaging two duplex arms, which is supported by single molecule studies. Biochemical analyses verify that MHF preferentially engage DNA forks or various four-way junctions independent of the junction-site structure. Genetic experiments provide evidence that the observed DNA-binding interface of MHF is important for cellular resistance to DNA damage. These results provide insights into how the MHF complex recognizes branched DNA and stimulates FANCM activity at such a structure to promote genome maintenance.

Introduction

Fanconi anemia (FA) is a genetic disease characterized by congenital developmental defects, progressive bone marrow failure, and early incidence of cancers ¹. Substantial evidence has linked the FA pathway to the repair of DNA interstrand crosslinks (ICLs) ². Sixteen FA complementation groups and their corresponding genes have been identified to date ^{3,4}.

[†]To whom correspondence should be addressed. patrick.sung@yale.edu (P.S.), yong.xiong@yale.edu (Y.X.).

*These authors contributed equally to this work

Author Contributions. Q.Z. and Y.X. determined the DNA-MHF crystal structure; D.S. performed the *in vitro* mutagenesis, protein purifications and biochemical experiments; A.S. and Y.X. determined the Apo-MHF structure; T.R.S. performed the cellular experiments; D.S., A.M. and Q.Z. performed the single molecule experiments; Q.Z., A.S., X.Z. and M.T. contributed to protein preparation; S.M. and L.R. supervised the single molecule experiments; A.R.M. supervised the cellular experiments; P.S. and Y.X. supervised the research; Q.Z., D.S., P.S. and Y.X. wrote the manuscript.

Conflict of interests. The authors declared no conflict of interests.

Data deposition. The atomic coordinates and structure factors for the apo-MHF1-MHF2 and MHF1-MHF2-DNA structure have been deposited in the Protein Data Bank with accession codes: 4NE6 (MHF-Se), 4NE5 (MHF-apo1), 4NE3 (MHF-apo2), 4NE1 (MHF-DNA2), and 4NDY (MHF-DNA1).

Several FA proteins (FANCA, -B, -C, -E, -F, -G, -L and -M) and their associated factors form the FA nuclear core complex that mediates the monoubiquitination of the FANCI-FANCD2 (ID) complex leading to the activation of the FA pathway⁵.

FANCM is a crucial member of the FA core complex. Orthologs of FANCM exist in archaeobacteria (Hef) and yeast (Mph1 and Fml1)^{6, 7, 8, 9}. FANCM and its orthologs possess a DNA-dependent ATPase activity capable of efficiently processing DNA intermediates, e.g. the D-loop and Holliday Junction (HJ) arising from homologous recombination events, and mediating the reversal of DNA replication fork structures^{10, 11}. These attributes of FANCM and its orthologs are likely important for homologous recombination regulation and replication fork repair^{9, 12}.

FANCM associates with a pair of histone-fold proteins called MHF1 and MHF2^{13, 14}, which form a tetramer harboring two copies of the heterodimer^{15, 16}. Cells depleted of either MHF protein recapitulate the FANCM sensitivity profile to various DNA damaging agents and are impaired for FA pathway activation^{13, 14}. The MHF complex (also known as CENP-S and CENP-X) also possesses centromere-specific functions¹⁶ and is responsible for targeting FANCM to these sites¹⁵. FANCM-MHF has been suggested to facilitate the replication of centromeric DNA. Biochemically, FANCM prefers to bind branched DNA structures such as model Holliday junctions and replication forks¹¹. The MHF complex also binds DNA and enhances the DNA replication fork reversal and HJ branch migration activities of FANCM^{13, 14}. It has been suggested that MHF confers to FANCM a higher degree of specificity for branched DNAs, which are common intermediates in homologous recombination and replication fork regression^{13, 14}.

How the MHF complex differentiates branched DNA from linear DNA is unknown, even though the crystal structures of MHF from different species^{15, 16, 17} have shown a tetrameric histone-fold arrangement with a potential DNA binding patch. To understand the mechanism by which MHF senses branched DNA, we used a combination of X-ray crystallographic, biochemical, single molecule, and *in vivo* functional studies to investigate the interaction between human MHF and various forms of DNA. Our results support a model wherein MHF recognizes a pair of DNA duplex arms to stimulate FANCM activity at branched DNA.

Results

Crystal structures of apo-MHF and its complex with DNA

To illuminate the MHF-DNA interaction, we determined the structures of human MHF in three crystal forms and a 26 bp dsDNA-MHF complex in two crystal forms (Fig. 1a, Supplementary Figure S1, Table 1). The MHF apo-structures diffracted to resolutions of 2.5–1.8 Å, with tetrameric MHF structures nearly identical to previously published apo MHF structures (Ca root-mean-square deviation, RMSD 0.5–1.1 Å)^{15, 16, 17}. The two MHF-DNA complex crystals, named MHF-DNA1 and MHF-DNA2, diffracted to resolutions of 7.2 and 6.5 Å, respectively. MHF-DNA2 is related to MHF-DNA1 by doubling the unit cell dimension in the C-axis due to a translational non-crystallographic symmetry. Apart from this translational doubling of the unit cell content, the molecules in the two crystals have virtually identical conformations and packing environment. For clarity, we use MHF-DNA1 to describe the structure. The asymmetric unit of the crystal contains two DNA duplexes and five MHF tetramers (labeled 1–5) in three interaction scenarios: MHF tetramer 1 contacts two DNA duplexes, one on each side of the tetramer (double-side binding); each of the MHF tetramers 2 and 3 interacts with one DNA duplex on one side of the tetramer (single-side binding); the MHF tetramers 4 and 5 do not interact with DNA.

The excess MHF molecules in the crystal are likely due to the length of the dsDNA allowing for the binding of multiple MHFs and also due to crystal packing.

The MHF-DNA complex structure has a very defined conformation with excellent electron density, in which the secondary structural elements of MHF and DNA major/minor grooves are clearly discernable (Fig. 1b, Supplementary Figure S2). Although side chain information cannot be directly revealed at 6.5–7 Å resolution, the well-defined secondary structure elements allow for accurate estimate of the side chain locations using the apo MHF structures as reference. The MHF tetramers in the DNA complex crystal adopt similar conformations to the published apo MHF structures (Supplementary Figure S1d), with slight structural variations (C α RMSD 2.1–2.7 Å) possibly due to DNA-binding perturbations and/or intrinsic positional uncertainties at 6.5–7 Å resolution. It is likely there are local differences at the DNA-binding interface of MHF with or without bound DNA, however, at our resolution we cannot resolve such detailed differences. The quality and correctness of the structure determination are demonstrated by the unambiguous simulated annealing omit map (Fig. 1b, S2), unbiased electron density for DNA at positions without models (Supplementary Figure S2), as well as good refinement statistics (R/R-free of 26.8%/28.8% for MHF-DNA1, Table 1). In addition, unbiased difference electron density for residues 107–118 in the previously unobserved C-terminus of MHF1 is clearly visible without any prior model information (Fig. 1b) (discussed below), further validating the structure.

The DNA-binding interface on MHF

A conserved MHF-DNA interaction is observed for both the single- and double-sided binding scenarios in the crystals. A superposition of all the independent MHF-DNA interfaces clearly reveals the conserved DNA-binding mode of MHF (Fig. 1a inset). The MHF-DNA interface buries 1040 Å² surface area per interaction site. To simplify the description, we divide the positively charged circumference of the MHF tetramer into top (K73, R74, K94, K99 and R110 of MHF1), shoulder (R18 of MHF1 and R17, R11, K12, K27, K29 and H20 of MHF2), arm (MHF1 C-terminus), and bottom (K44 of MHF1, R64 of MHF2) patches (Fig. 2a). The MHF residues at the DNA-binding interface are conserved among many vertebrates including human, chimpanzee, mouse, bird, and fish (Supplementary Figure S3a, b).

The MHF-DNA structure shows that residues in the top, shoulder, and arm patches of MHF are close to the DNA backbone, while the bottom patch is situated away from the binding interface (Fig. 2a). Although the MHF-DNA interface covers both the minor and the major grooves of DNA, MHF primarily interacts with DNA sugar-phosphate backbones and there does not appear to be base specific interactions in the grooves. This backbone-only recognition is a common feature of nonspecific DNA binding proteins. To validate the structural observation, we systematically investigated the charged MHF residues involved in DNA binding (Fig. 2b and c). As predicted from the structure, mutations of the residues in the shoulder and top patches abolish MHF binding to HJ, forked and duplex DNAs¹⁴, while mutations in the bottom patch have little effect on DNA binding (Fig. 2b and c).

The positively charged C-terminal region of MHF1, which is disordered in all apo MHF structures, also engages in interactions with DNA and becomes structured. Extending from the end of the structured apo-MHF1 C-terminal α -helix, residues 107–118 form approximately two more α -helical turns followed by a loop region. These newly structured residues are located next to the negative charged DNA phosphate backbone (Fig. 1b). Although we cannot accurately model side chains at the low resolution, it is evident that the positive-charged clusters (including R110, K111, K114, K115, K116, K117 of MHF1) in this arm region are positioned close to the DNA backbone forming potential electrostatic interactions (Fig. 1b). The MHF1 C-terminus is only observed at the sites where the

molecule interacts with DNA, while it remains disordered in those MHF1 copies in the crystal that do not contact DNA. This is consistent with and provides a structural explanation for the observation that the C-terminal region of MHF1 is necessary for DNA binding¹³.

DNA condensation by binding to both sides of MHF

In the crystal structure, MHF tetramer 1 (Fig. 1a) engages a pair of DNA duplexes through the positively charged patches on the side and upper circumference of the MHF tetramer as well as the MHF1 C-terminal arm (Fig. 2a, Fig. 3a). Interestingly, the two DNA helices are not co-planar, but are oriented in a stacked conformation with the helical paths of the duplexes closely packing against each other (Fig. 3a). A potential consequence of DNA binding by both sides of the MHF tetramer is the condensation of long DNA through bending and looping. Indeed, dsDNA condensing by MHF has been observed by electron microscopy¹⁴.

We conducted single-molecule experiments to further verify the condensation of DNA by MHF. We used the single molecule technique of Tethered Particle Flow Stretching (TPFS)¹⁸ to show that MHF association indeed induces the shortening of dsDNA (Fig. 3b and c). In the TPFS experiment, one end of a dsDNA (4,200 bp) was fixed and the other end was tethered to an optically visible bead, whose position was recorded throughout the experiment. The flexible DNA was stretched to an extended conformation by a calibrated constant flow force of ~0.5 pN (Supplementary Figure S4). The introduction of MHF led to a significant contraction of the DNA chain by approximately a third of its length (~1400 bps) in our experimental setup (Fig. 3c). The DNA condensation was indeed caused by MHF binding, as revealed by testing MHF mutants (Fig. 3c). Specifically, a shoulder patch mutant of MHF (R18A of MHF1 and R11A/R17A of MHF2) deficient in DNA binding failed to condense the DNA, while a bottom patch mutant that retains DNA binding (K44A of MHF1 and R64A of MHF2) was proficient in this regard (Fig. 3c).

MHF and MHF-FANCM complex prefer branched DNA over dsDNA

The architecture of the MHF-DNA interaction (Fig. 2a) strongly suggests an ability of MHF to specifically engage branched DNAs¹⁴, as both sides of MHF are available for simultaneous DNA binding. To assess the relative affinity of the MHF complex for branched versus linear dsDNA, we performed electrophoretic mobility shift assays (EMSA) with MHF and three different substrates: dsDNA, DNA fork, and Holliday junction (HJ) (Fig. 4a and b). The results revealed the order of MHF binding preference as: HJ > DNA fork > dsDNA. We then performed competition assays in which a nucleoprotein complex of MHF with radiolabeled dsDNA was challenged with either cold dsDNA or HJ (Fig. 4c). The results clearly showed that HJ is a much better competitor than dsDNA (Fig. 4d). At 30 nM cold competitor concentration, which was the same as that of the radiolabeled dsDNA, there was a ~5 fold increase in the amount of released radiolabeled DNA when challenged with HJ versus dsDNA (Fig. 4d).

As MHF interacts with FANCM and stimulates its DNA branch migration activity^{13, 14}, we investigated whether the presence of MHF-interacting FANCM fragment (FANCM-F) changes the DNA binding properties of MHF. We performed the same competition experiments with the complex of MHF and the FANCM fragment (spanning residues 661 to 800) that harbors the MHF1 interaction domain¹⁴. While this fragment of FANCM does not bind DNA on its own (Supplementary Figure S5a), the FANCM-F/MHF complex, when compared to MHF alone, has an enhanced DNA binding ability (Fig. 4a and b; Supplementary Figure S5b). FANCM-F/MHF shows the same DNA binding preference as MHF, i.e. HJ > DNA fork > dsDNA (Fig. 4a and b).

MHF preference to branched DNA with adjacent duplexes

We notice that the MHF tetramer 1 orients the pair of bound DNA duplexes at a $\sim 55^\circ$ angle (Fig. 5a), which resembles the angle between the two arms of the four way junction DNA, three way DNA junction, and paused DNA fork in the “stacked X” conformation previously observed in crystal structures^{19, 20, 21}. The observed angle is also close to the angle of $\sim 60^\circ$ of the HJ and three-way junction in solution^{22, 23}. This broad geometry resemblance may explain the preference of MHF for branched DNA substrates, where two adjacent duplexes on one side of the junction are bound simultaneously by one MHF tetramer. This would provide a general mechanism by which MHF engages a pair of crossover duplexes in branched DNA structures regardless of the overall architecture of the DNA.

To verify that engagement of duplex arms by MHF is independent of the architecture of the junction point, we assembled HJ-like substrates with a 5, 10 or 20-nucleotide long unpaired bubble region at the junction site (Fig 5b). EMSA experiments with these HJ-like substrates and control linear duplexes with the same size of DNA bubble revealed that the branched HJ-like DNA remains the preferred substrate in each case (Fig. 5c, d, e and S6). As observed before (Fig 4a and b), the presence of FANCM-F fragment increases DNA binding affinity but does not change substrate preference (Fig. 5c, d, e and S6). These data provide additional support to the premise that the preference of MHF and MHF-FANCM for branched DNA structures stems from their ability to simultaneously engage two crossed duplex arms.

The MHF-DNA binding interface is functionally important

It has been reported that the interaction of MHF with FANCM is important for the upregulation of the DNA branch migration activity of FANCM^{13, 14}. To test the importance of the observed MHF-DNA interaction in FANCM stimulation, we performed branch migration assays with FANCM alone or together with either wild type (WT) or various mutants of MHF. We used in these reactions a radiolabeled model movable replication fork (MRF) as the substrate^{11, 13}. While MHF enhanced the branch migration activity of FANCM, the top-and shoulder patch mutants that are deficient in DNA binding (R18A/R11A/R17A, K73A/K94A/K99A/R110A, K73A/R74A, K12A/H20A/K27A/K29A) failed to do so. Consistently, the bottom-patch MHF mutant that is DNA binding proficient (K44A/R64A) stimulated FANCM normally (Fig. 6a, b, c, d). Importantly, all the mutants preserve the WT oligomerization state (Supplementary Figure S3c) and interact with FANCM as well as wild-type MHF (Fig. 6e, f, g, h), suggesting the overall structure is not significantly perturbed by any of these mutations. These results demonstrate that the FANCM stimulatory effect of MHF requires its DNA binding activity.

We also validated the biological relevance of the MHF DNA-binding attribute using cell-based assays. For this, we depleted MHF1 endogenous to HeLa cells via shRNA treatment (Supplementary Figure S7a) and then examined the ability of ectopically expressed, shRNA-resistant wild type or mutant MHF1 to complement the phenotypes stemming from MHF1 ablation. By co-immunoprecipitation, we verified that the DNA-binding deficient mutant (K73A/K94A/K99A/R110A of MHF1) is incorporated into the higher order complex with MHF2, FANCM, and FAAP24 (Fig. 7a). However, cells expressing the mutant protein remained sensitive to camptothecin (CPT) and mitomycin C (MMC) (Fig. 7c). Consistent with these results, the mutant MHF1 failed to support FANCD2 monoubiquitination (Fig. 7b) and was also defective for chromatin co-localization of FA core complex members (Fig. 7d) upon MMC treatment. We note that previous work^{13, 14, 15} also showed that mutations in the top patch residues (K73, R74 of MHF1) render the MHF complex impaired in DNA binding, chromatin co-localization of FA proteins, and FANCD2 monoubiquitination at

cellular level. Therefore, the DNA binding interface observed in the MHF complex is clearly important for the functional integrity of the FA pathway of DNA repair.

Discussion

MHF plays an important role in facilitating the processing of multiple types of branched DNAs in the FA pathway of DNA damage repair. Our X-ray crystallographic studies on human MHF provide evidence that MHF recognizes two crossover DNA duplexes, a common feature in various DNA junction structures that occur during DNA replication and repair in cells. This attribute of MHF would enable MHF to recognize a wide range of branched DNA junction structures of different architectures, such as forked and Holliday Junction DNAs.

The preference of MHF for branched DNA suggests a potential mechanism by which this protein complex facilitates the localization of FANCM to DNA junctions. We have analyzed possible protein-DNA interactions at a DNA junction using structural information from our MHF-DNA complex, the MHF-FANCM complex¹⁵, an ideal HJ DNA, and a homology-modeled FANCM helicase domain based on the crystal structure of Hef²⁴ (Fig. 8a). According to our model, simultaneous interactions of MHF with two DNA arms and FANCM promote the recruitment of FANCM to the DNA junction (Fig. 8a, b). The MHF-FANCM interaction could further help position the helicase domain of FANCM at the junction crossover point to mediate DNA branch migration, as the distance between the MHF-binding site and the DNA junction appears to match the dimension of the homology-modeled helicase domain (Fig. 8a). As a result, the assembly of the MHF-FANCM complex on branched DNA is expected to minimize unproductive association of FANCM with duplex DNA away from the junction, and to enhance the efficiency of DNA branch migration by orienting the FANCM active site in the optimal geometry. Considering MHF's biochemical role in stimulating the translocase activity of FANCM, such resemblance brings new possibility into MHF's working mechanism. MHF might not only increase the FANCM binding affinity to branch DNA substrate, but also confer some dynamic property to make FANCM-MHF complex a fully functional translocase targeting branched DNA site.

The results presented herein provide a framework for understanding the mechanisms by which the conserved MHF complex engages various branched DNA structures and facilitates the FA pathway of genome maintenance and replication fork repair. Insight into the clinical impact of MHF is currently limited by the relatively rare occurrence of MHF mutations in patients. There are only 19 mutations, mostly missense substitutions, in MHF1 gene out of 7741 unique samples in the Catalogue Of Somatic Mutations In Cancer (COSMIC) database (<http://www.sanger.ac.uk/genetics/CGP/cosmic/>). Cancer linkage to MHF1 or MHF2 has not been identified. However, because of the important cellular functions of MHF complex in FA pathway activation and DNA genome maintenance, it is possible that FA patients with mutations in MHF1 and MHF2 will be discovered in the future.

Methods

Protein overexpression and purification

The human MHF complex was expressed as described¹³. Briefly, Rosetta *E. coli* BL21 (DE3) cells (Novagen) were transformed with a duet vector expressing both MHF1 (GST-tagged) and MHF2 (6XHis-tagged) proteins. After cell lysis, the MHF complex was purified by a combination of Talon (Clontech), glutathione affinity (GE Healthcare), treatment with TEV (Tobacco Etch Virus) protease to cleave off the GST tag on MHF1, hydroxyapatite (GE Healthcare), Mono Q (GE Healthcare), and Superdex S200 gel filtration

chromatography (GE Healthcare). The MHF mutants used in this study were generated using the QuickChange Site-Directed Mutagenesis kit (Stratagene), verified by DNA sequencing, expressed, and purified using the same procedure. SeMet-containing MHF complex was expressed in B834 competent cells (Novagen) and grown in M9 minimal media (Sigma-Aldrich) containing Se-Met (Sigma-Aldrich), and purified using the same procedure.

MHF-DNA complex was obtained by mixing MHF tetramer with a DNA that enables the formation of duplex DNAs from the self-complementary sequence 5'-ATCAATATCCACCGGTGGATATTGAT-3' (Integrated DNA Technologies), in 2:1 molar-ratio and further purified by gel-filtration chromatography on a Superdex 200 column buffered with 20 mM Tris pH 8.0, 150 mM NaCl, and 0.1 mM TCEP.

The FANCM-F fragment (661-800) was cloned in the pMAL-TEV (New England Biolabs) vector and expressed in Rosetta cells with an N-terminal MBP (maltose binding protein) tag. The protein was purified using amylose agarose (GE Healthcare), Mono Q and gel-filtration in Superdex 200 column. The FANCM-F/MHF1-MHF2 complex was purified as follows. Cells expressing MBP-tagged FANCM-F and GST-tagged MHF1-MHF2 were mixed and an extract was prepared. Purification was achieved by glutathione and amylose affinity steps. TEV protease was mixed with protein eluted from the amylose resin (1:1000 molar ratio) followed by an overnight incubation. The cleaved tags were removed by gel filtration in Superdex 200.

Crystallization and structure determination

The apo MHF was crystallized using the macro-batch under oil method²⁵. A mixture of protein and reservoir buffer was deposited under a layer of 2:1 paraffin (Fluka): silicon (Hampton Research) oil and allowed to equilibrate at room temperature. Crystals were obtained in crystallization buffers containing (a) 0.1M sodium phosphate pH 6.5 with 12% PEG 8000, (b) 0.1M HEPES pH 7.0 with 20% PEG8000, and (c) 0.1M magnesium chloride with 0.1M HEPES pH 7 and 15% PEG4000. The SeMet-containing MHF complex was crystallized under the same conditions. Crystals were frozen using 30% glycerol as a cryoprotectant. Diffraction data were collected at the Advanced Photon Source beamline 24-ID and the National Synchrotron Light Source beamline X29A. Data were processed and scaled using HKL2000²⁶ (Table 1). The structure was solved by a combination of SeMet SAD and MAD phasing at 2.1 Å resolution using Shelx²⁷ and Solve²⁸. A 2-fold non-crystallographic symmetry (NCS) averaging using *DM*²⁹ significantly improved the electron density map, which allowed Buccaneer³⁰ for automatic model building of 308 out of 438 total ASU residues. The model was subsequently refined using *Phenix*³¹ and manually rebuilt in *Coot*³² to a final R/R_{free} of 0.230/0.262. The structures in the other two crystal forms were solved by molecular replacement using the model obtained for the SeMet data and refined to R/R_{free} of 0.223/0.264 (MHF-apo1) and 0.231/0.252 (MHF-apo2), respectively. Data and refinement statistics are summarized in Table 1.

MHF-DNA crystals were grown in 0.5–0.7M ammonium phosphate at room temperature using the macro-batch under oil method, and frozen in liquid nitrogen with 15% ethylene glycol added as cryo-protectant. Diffraction data were collected at 100K on beamline 24-IDC at the Advanced Photon Source. Data were processed using HKL2000 (MHF-DNA1 in Table 1) to 7.2 Å ($I/\sigma \sim 1$, $CC_{1/2} \sim 0.14$ by the Pearson correlation coefficient method³³) in space group P321. Molecular replacement (MR) was performed in PHASER^{29, 34} using the MHF heterodimer as a search model. This gave an initial solution of three heterodimers (Supplementary Figure S2a). Further MR using a 13bp ideal B-DNA model (generated by Coot with the sequence 5'-AAAAAAAAAAAAA-3' in one strand and the complementary sequence in the other strand) yielded two solutions in a slightly misaligned head-to-tail

geometry (Supplementary Figure S2b, c). Subsequent refinement and density modification revealed a continuous double helical density, accommodating a 26bp ideal B-DNA, at a position different from that of the initial 13bp DNA model, confirming the electron density for the DNA is not model biased (Supplementary Figure S2c, d). Further refinement and additional real-space MR search in the electron density located the last two MHF heterodimers in the crystal. Refinement was carried out using Phenix³¹ and Refmac³⁵ using local restraints from a reference structure of apo-MHF³⁶, jelly-body³⁷, and NCS³⁷ restraints, yielding a refined model with R/R_{free} of 32.3%/34.4% and good stereochemical geometry.

A second, related crystal form of the MHF-DNA complex (MHF-DNA2) was obtained, which diffracted to 6.5 Å resolution ($I/\sigma \sim 0.5$, $CC_{1/2} \sim 0.16$ by the Pearson correlation coefficient method³³). The two crystals have very similar cell dimensions in *a* and *b* directions but the cell constant in the *c* direction is doubled in MHF-DNA2. Inspection of the native Patterson map revealed a strong NCS translational peak of ~80% of the origin peak height at the fractional coordinates of (0, 0, 0.5). This is consistent with that the unit cell of MHF-DNA2 consists of two translationally related layers of molecules along the *c*-axis. The molecules in each layer have virtually identical structures and packing environment as those in MHF-DNA1. The solution obtained in MHF-DNA1 was used as a search model to solve the structure of MHF-DNA2, which is subsequently refined to R/R_{free} of 30.7%/37.9%.

A thorough inspection of the diffraction data revealed that both MHF-DNA complex crystals are twinned, with twinning fractions close to 0.5 (Table 1) in the P3 space group resulting in an apparent space group of P321. Interestingly, the conventional twinning tests, such as Wilson ratios and cumulative intensity distribution, did not detect twinning. This is presumably due to the high degree of NCS, especially the translational NCS in MHF-DNA2, which may alter the apparent intensity distribution. Nonetheless, the L test³⁸ and the maximum likelihood method implemented in Phenix Xtriage³¹ clearly revealed twinning with the twinning operator “*k, h, -l*” (Supplementary Figure S8). Diffraction data for both MHF-DNA1 and MHF-DNA2 were reprocessed in P3 space group. Care was taken to maintain a consistent and independent set of R-free flags, with those from the P321 data kept and expanded based on the twinning operator using *Phenix.reflection_tools*³¹.

The amplitude based twin refinement in *Refmac5*³⁷ resulted in a substantial decrease of R_{free} factors of ~6% in both MHF-DNA crystal forms (R/R_{free}: 26.8%/28.8% for MHF-DNA1 and 28.3%/31.3% for MHF-DNA2). The slightly higher R factors of the MHF-DNA2 structure are probably due to the strong translational NCS in the crystal, as pseudo-translation is known to introduce artificially high R factors due to intensity modulations^{39, 40}. The B-factors of the refined structures are consistent with the Wilson B factors calculated from the diffraction data (Table 1; Table S2; Supplementary Figure S1b, c), although Wilson statistics at low resolution are not accurate. The B-factors for the DNA and the protein components are similar, indicating similar ordering of these components in the crystals. More importantly, the twin refinement significantly improved the resulting electron density maps, exemplified by the appearance of unbiased difference electron density for the MHF1 C-terminal region without any model information (Fig. 1b; Supplementary Figure S2g), which only appeared after successful twin refinement.

Radiolabelled DNA substrate preparation

Radiolabeled ds, Fork, HJ and MRF substrates were prepared from synthetic oligos (Supplementary Table S1). The DNA hybridization reactions were conducted by heating equimolar amounts of the indicated oligonucleotide mixture at 95°C for 10 min in buffer H (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 100 mM NaCl) followed by slow cooling to room

temperature. DNA substrates were separated from un-annealed oligonucleotides in 10% non-denaturing polyacrylamide gels run in TAE buffer (40 mM Tris acetate, pH 7.4, 0.5 mM EDTA) and were recovered from the gel by electroelution in dialysis tubing in TAE buffer.

DNA binding assays

Wild type and mutant MHF complexes were incubated with various radiolabeled DNA substrates (double-stranded DNA (dsDNA), movable replication fork (MRF), and Holliday junction (HJ)). The reaction mixtures were analyzed by mixing protein with various concentration of DNA substrates. The concentration of MHF complex is calculated as molar concentration of the tetramer. Reactions were conducted in triplicates to obtain the mean and the standard deviation values. For the DNA binding competition assays, MHF or the FANCM-F/MHF complex was incubated with a mixture of ds-, forked, and HJ DNAs at 30 nM molar concentration each in buffer (25 mM Tris (pH 7.5), 1 mM DTT, 100 µg/mL bovine serum albumin) containing 150 mM KCl for 10 min at 37°C. After addition of 5 µl loading buffer (50% glycerol, 20 mM Tris-HCl (pH 7.4), 2 mM EDTA, 0.05% orange G), the reaction mixtures were resolved in 6.5% native polyacrylamide gel in TBE buffer (45 mM Tris-borate, 1mM EDTA (pH 8.0)) at 4°C. Gels were dried onto Whatman DE81 paper (Whatman International Limited) and then analyzed in a Personal Molecular Imager FX PhosphorImager (Bio-Rad). Reactions were conducted in triplicates to obtain the mean and the standard deviation values.

Cold DNA Competition assays

For competition, the MHF complex (100 nM) was preincubated with 30 nM radiolabeled dsDNA substrate for 5 min at 37°C. Subsequently, the indicated concentration of either unlabeled dsDNA or HJ DNA was added, followed by a 5-min incubation at 37 °C. The reaction mixtures were resolved and analyzed as above. Reactions were conducted in triplicates to obtain the mean and the standard deviation values.

DNA branch migration assay

The reactions (10 µl) for branch migration assays with MRF contained 10 nM protein with 5 nM DNA substrate in the presence of ATP. Reactions were carried out at 37 °C for the indicated periods. The resulting mixture were deproteinized for 20 min at 37 °C with 2 mg/ml Proteinase K and 0.4% SDS and resolved by native PAGE through 8% polyacrylamide gels in TBE. Assays were conducted in triplicates to obtain the mean and the standard deviation values.

MHF and FANCM-F Pull-down assay

10 µg of wild type MHF or MHF mutant (K73A/R74A, K12A/H20A/K27A/K29A, K73A/K94A/K99A/R110A, K44A/R64A, or R18A/R11A/R17A) was incubated with the same molar ratio of MBP-His-FANCM-F (FANCM fragment 661-800) and incubated in 30 µl of buffer (25 mM Tris-HCl at pH 7.5, 0.01% NP-40, 50 mM KCl, 1 mM DTT) for 1 hour at 4 °C. Amylose agarose resin (10 µl; New England Biolabs) was added, followed by gentle mixing for 30 min at 4°C. After washing the beads twice with 200 µl of buffer, bound proteins were eluted with 30 µl of 2% SDS. 15% of the supernatant and elution fractions, and 2% of total wash fraction were analyzed by 15% SDS-PAGE followed by Coomassie blue staining. Control experiments were conducted to show that the MHF mutants were not retained on the Amylose resin without the inclusion of the MBP-tagged FANCM-F.

Homology modeling of the FANCM helicase domain

A crystal structure (PDBID: 1WP9) is available for *Pyrococcus furiosus* Hef²⁴, the FANCM ortholog. Hef has 764 amino acid residues and 30/70% sequence identity/similarity to the corresponding helicase and C-terminal regions of FANCM. Because of the high sequence similarity, we were able to generate a homology model of the FANCM helicase domain based on the Hef crystal structure²⁴ using the automated modeling server I-TASSER⁴¹, which has consistently ranked as the top performer in the CASP (Critical Assessment of Techniques for Protein Structure Prediction) modeling contests.

Tethered Particle Flow Stretching (TPFS)

WT or mutant MHF was diluted to 10 μ M in PBS with 1 mg/mL casein and 0.1% TWEEN and flowed into the cell a flow cell with a depth of 0.5 mm and a width of 5 mm. Measurements were taken at the center of the flow cell. To stretch the DNA, beads were tracked under constant flow force of \sim 0.5 pN generated using a syringe pump (OEM syringe pump, Harvard Apparatus). The flow force was calibrated by measuring the mean displacement of the bead at various flow rates and then fitting to an experimentally calibrated wormlike chain model that relates the extension of the DNA-tethered beads with the measured force exerted by optic tweezers⁴⁴. To obtain the flow force of \sim 0.5 pN, we first calculated the extension of the tethered bead predicted from the wormlike chain model⁴⁴, and subsequently tuned the flow rate to match the experimental extension (Supplementary Figure S4a) to the theoretical prediction (Supplementary Figure S4b). Detailed information regarding instrumental setup and calibration is in Supplementary Methods.

The flow rate equivalent to \sim 0.5 pN force is 64 μ L/min in our experimental setup, which was used for all TPFS assays on both WT MHF and MHF mutants. The measurements were repeated for a minimum of 10 tethers in each case. As the measured total extension includes both the length of the DNA and the size/radius of the tethered bead, the apparent initial extensions can vary up to 20% from bead to bead due to the intrinsic variations in bead size. Additionally, the difference in bead size results in a fluctuation in flow force from bead to bead. Despite the variations, the same length reduction (corresponding to \sim 1400bp long) from initial extension has been observed consistently upon the addition of all MHF capable of DNA-binding, but not for the DNA-binding impaired MHF mutant, suggesting the same MHF-induced DNA condensation phenomenon. Initially when few MHF molecules are bound, they are potentially in the single-side binding mode and there is little condensation occurring right away. This is probably because the DNA-bending, double-side binding energy is not enough to overcome the energy required for DNA bending under the experimental flow force. When enough MHF proteins bind, the collective binding energy overcomes the energy to bend the DNA and the DNA begins to condense. The initial bending will then bring duplex regions of the DNA close together creating a geometry that facilitates the binding of the next MHF molecules. This continues until an equilibrium is reached in the experimental condition and condensation stops. This type of condensation has been observed with other DNA binding proteins such as histone H1⁴⁵.

Antibodies and drug sensitivity tests

Anti-Flag (1:500), -MHF1 (1:1000), -MHF2 (1:500), -FANCM (1:2000), -FAAP24 (1:3000) and -FANCD2 (1:4000) antibodies were prepared and drug sensitivity assays were performed. 200 HeLa cells transduced with lentiviruses carrying either control shRNA or shMHF1-UTR were seeded per 10 cm dish containing the indicated concentration of MMC or camptothecin (in Fig. 6c). After 10 days, cells were fixed, stained, and visible colonies were counted. Assays were conducted in triplicates to obtain the mean and the standard deviation values.

Co-immunoprecipitation of FANCM and MHF

HeLa cells were transduced with lentiviruses carrying shMHF1-UTR (5'-GATAATGTGTACTGCGTTA-3', throughout this study). The cells were subsequently transduced with retroviruses carrying empty vector (pMIEG3), Flag-WT-MHF1, or Flag-MHF1 mutant (K73A/K94A/K99A/R110A) and treated with 100 ng/mL MMC for 16h. Cell lysate was prepared and treated with anti-FLAG resin (Sigma-Aldrich). Immunoprecipitated protein complexes were analyzed by immunoblotting with the corresponding antibodies (anti-FLAG, -MHF2, -FAAP24, -FANCM).

Chromatin localization assay

HeLa cells were transduced with lentiviruses carrying either control shRNA (SCB: 5'-CCTAAGGTTAAGTCGCCCTCG-3', throughout this study) or shMHF1-UTR. Then, HeLa-MHF1-UTR cells were transduced with retroviruses carrying either empty vector, shRNA resistant WT-MHF1 cDNA or mutant-MHF (K73A/K94A/K99A/R110A). Cells were treated with 100ng/ml MMC for 16h, followed by cellular fractionation into total lysate and chromatin fraction. Cell-equivalent volumes from total cell lysate and chromatin fraction were separated by SDS-PAGE and immunoblotted with antibodies (Supplementary Figure S9a).

FANCD2 activation assay

HeLa cells were transduced with lentiviruses carrying either control shRNA or shMHF1-UTR. Subsequently, cells were transduced with retroviruses carrying empty vector (pMIEG3), shRNA resistant WT-MHF1 cDNA, or mutant-MHF (K73A/K94A/K99A/R110A). EGFP positive cells were sorted by fluorescence-activated cell sorting (FACS). Cells were treated with 100 ng/ml MMC for 16h, lysed, and immunoblotted with anti-FANCD2 antibody(1:4000) (Supplementary Figure S9b).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the staff at the Advanced Photon Source beamline 24-ID, the National Synchrotron Light Source beamlines X25 and X29A. This work was supported by an Alex's Lemonade Stand Foundation Innovation Award and a March of Dimes Foundation Basil O'Connor Starter Scholar Research award (Y.X.), and by grants from the US National Institutes of Health (HL084082, GM57814, ES015632, CA168635 and CA92584).

References

1. Auerbach AD. Fanconi anemia and its diagnosis. *Mutat Res.* 2009; 668:4–10. [PubMed: 19622403]
2. Knipscheer P, et al. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science.* 2009; 326:1698–1701. [PubMed: 19965384]
3. Kashiyama K, 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. [PubMed: 23623389]
4. Kim H, D'Andrea AD. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes Dev.* 2012; 26:1393–1408. [PubMed: 22751496]
5. Kottemann MC, Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature.* 2013; 493:356–363. [PubMed: 23325218]
6. Meetei AR, 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. [PubMed: 16116422]

7. Sun W, 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. [PubMed: 18851838]
8. Prakash R, et al. Yeast Mph1 helicase dissociates Rad51-made D-loops: implications for crossover control in mitotic recombination. *Genes Dev*. 2009; 23:67–79. [PubMed: 19136626]
9. Singh TR, et al. Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. *Blood*. 2009; 114:174–180. [PubMed: 19423727]
10. 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 U S A*. 2008; 105:16107–16112. [PubMed: 18843105]
11. 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. [PubMed: 18206976]
12. 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. [PubMed: 19465393]
13. Singh TR, et al. MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM. *Mol Cell*. 2010; 37:879–886. [PubMed: 20347429]
14. Yan Z, et al. A histone-fold complex and FANCM form a conserved DNA-remodeling complex to maintain genome stability. *Mol Cell*. 2010; 37:865–878. [PubMed: 20347428]
15. Tao Y, et al. The structure of the FANCM-MHF complex reveals physical features for functional assembly. *Nat Commun*. 2012; 3:782. [PubMed: 22510687]
16. Nishino T, et al. CENP-T-W-S-X Forms a Unique Centromeric Chromatin Structure with a Histone-like Fold. *Cell*. 2012; 148:487–501. [PubMed: 22304917]
17. Yang H, et al. *Saccharomyces Cerevisiae* MHF Complex Structurally Resembles the Histones (H3-H4)₂ Heterotetramer and Functions as a Heterotetramer. *Structure*. 2012; 20:364–370. [PubMed: 22325783]
18. Schlingman DJ, Mack AH, Mochrie SG, Regan L. A new method for the covalent attachment of DNA to a surface for single-molecule studies. *Colloids Surf B Biointerfaces*. 2010; 83:91–95. [PubMed: 21130613]
19. Nowakowski J, Shim PJ, Prasad GS, Stout CD, Joyce GF. Crystal structure of an 82-nucleotide RNA-DNA complex formed by the 10–23 DNA enzyme. *Nat Struct Biol*. 1999; 6:151–156. [PubMed: 10048927]
20. Nowakowski J, Shim PJ, Stout CD, Joyce GF. Alternative conformations of a nucleic acid four-way junction. *J Mol Biol*. 2000; 300:93–102. [PubMed: 10864501]
21. Khuu P, Ho PS. A rare nucleotide base tautomer in the structure of an asymmetric DNA junction. *Biochemistry*. 2009; 48:7824–7832. [PubMed: 19580331]
22. Duckett DR, Murchie AI, Diekmann S, von Kitzing E, Kemper B, Lilley DM. The structure of the Holliday junction, and its resolution. *Cell*. 1988; 55:79–89. [PubMed: 3167979]
23. Murchie AIH, Clegg RM, Vonkitzing E, Duckett DR, Diekmann S, Lilley DMJ. Fluorescence Energy-Transfer Shows That the 4-Way DNA Junction Is a Right-Handed Cross of Antiparallel Molecules. *Nature*. 1989; 341:763–766. [PubMed: 2797209]
24. Nishino T, Komori K, Tsuchiya D, Ishino Y, Morikawa K. Crystal structure and functional implications of *Pyrococcus furiosus* hef helicase domain involved in branched DNA processing. *Structure*. 2005; 13:143–153. [PubMed: 15642269]
25. Chayen NE. The role of oil in macromolecular crystallization. *Structure*. 1997; 5:1269–1274. [PubMed: 9351804]
26. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A*. 1997; 276:307–326.
27. Sheldrick GM. A short history of SHELX. *Acta Crystallogr A*. 2008; 64:112–122. [PubMed: 18156677]
28. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr*. 1999; 55:849–861. [PubMed: 10089316]

29. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr.* 1994; 50:760–763. [PubMed: 15299374]
30. Cowtan K, Main P. Miscellaneous algorithms for density modification. *Acta Crystallogr D Biol Crystallogr.* 1998; 54:487–493. [PubMed: 9761844]
31. Adams PD, et al. The Phenix software for automated determination of macromolecular structures. *Methods.* 2011; 55:94–106. [PubMed: 21821126]
32. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004; 60:2126–2132. [PubMed: 15572765]
33. Karpplus PA, Diederichs K. Linking crystallographic model and data quality. *Science.* 2012; 336:1030–1033. [PubMed: 22628654]
34. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr.* 2007; 40:658–674. [PubMed: 19461840]
35. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr.* 1997; 53:240–255. [PubMed: 15299926]
36. Nicholls RA, Long F, Murshudov GN. Low-resolution refinement tools in REFMAC5. *Acta Crystallogr D Biol Crystallogr.* 2012; 68:404–417. [PubMed: 22505260]
37. Murshudov GN, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr.* 2011; 67:355–367. [PubMed: 21460454]
38. Padilla JE, Yeates TO. A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. *Acta Crystallogr D Biol Crystallogr.* 2003; 59:1124–1130. [PubMed: 12832754]
39. Jani D, et al. Sus1, Cdc31, and the Sac3 CID region form a conserved interaction platform that promotes nuclear pore association and mRNA export. *Mol Cell.* 2009; 33:727–737. [PubMed: 19328066]
40. Chook YM, Lipscomb WN, Ke H. Detection and use of pseudo-translation in determination of protein structures. *Acta Crystallogr D Biol Crystallogr.* 1998; 54:822–827. [PubMed: 9757096]
41. Zhang Y. I-TASSER server for protein 3D structure prediction. *Bmc Bioinformatics.* 2008; 9
42. Pouget N, Dennis C, Turlan C, Grigoriev M, Chandler M, Salome L. Single-particle tracking for DNA tether length monitoring. *Nucleic Acids Res.* 2004; 32:e73. [PubMed: 15155821]
43. Segall DE, Nelson PC, Phillips R. Volume-exclusion effects in tethered-particle experiments: bead size matters. *Phys Rev Lett.* 2006; 96:088306. [PubMed: 16606235]
44. Wang MD, Yin H, Landick R, Gelles J, Block SM. Stretching DNA with optical tweezers. *Biophys J.* 1997; 72:1335–1346. [PubMed: 9138579]
45. Bloomfield VA. DNA condensation by multivalent cations. *Biopolymers.* 1997; 44:269–282. [PubMed: 9591479]
46. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics.* 2002; Chapter 2(Unit 2):3. [PubMed: 18792934]
47. Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol.* 27:221–224. [PubMed: 19854763]

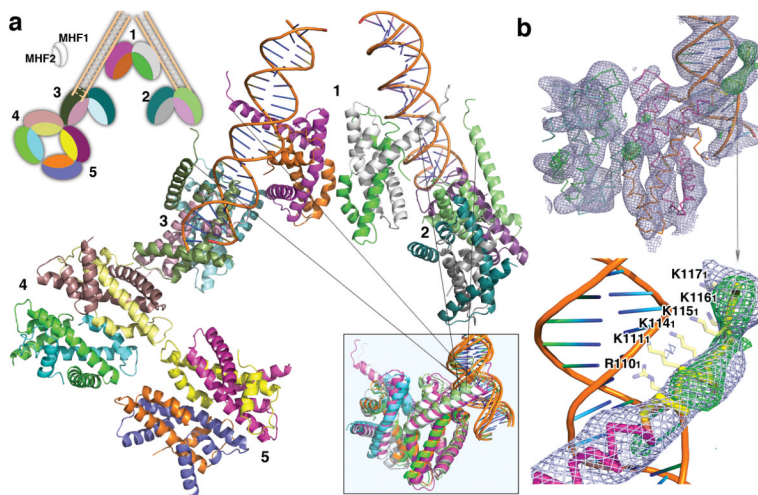


Fig. 1. Crystal structure of the MHF-DNA complex. **a**, The overall structure in the asymmetric unit of the MHF-DNA1 crystal. Two DNA duplexes (brown) and five MHF tetramers (labeled 1–5) are shown in ribbon representation. The top left inset shows the schematic view of the MHF complex (black-and-white ovals) and the organization of the components of the MHF-DNA complex (colored). The bottom right inset shows the superposition of the four observed MHF-DNA interaction interfaces. MHF tetramers 1 (two sites), 2, and 3 with associated DNAs are used in this comparison and the arrows point to their positions in the MHF-DNA complex. **b**, Unbiased electron density of the MHF-DNA interaction site. Simulated annealing omit map (blue mesh, $2mF_o-DF_c$ contoured at 1.1σ , 5 \AA around the molecules) was calculated with the displayed MHF tetramer 3 (wire representation) and DNA region (ribbon representation) omitted in the MHF-DNA2 structure. mF_o-DF_c difference density (green mesh, 2.6σ , 5 \AA around the molecules) was generated before any of the C-terminus (yellow) was built. Top panel, the overall MHF tetramer-DNA interaction site. Bottom panel, a close-up view of the unbiased electron density around the MHF1 C-terminal region (4 \AA around MHF1). The side chains of the positively charged residues (semi-transparent sticks) are shown in preferred rotamer positions.

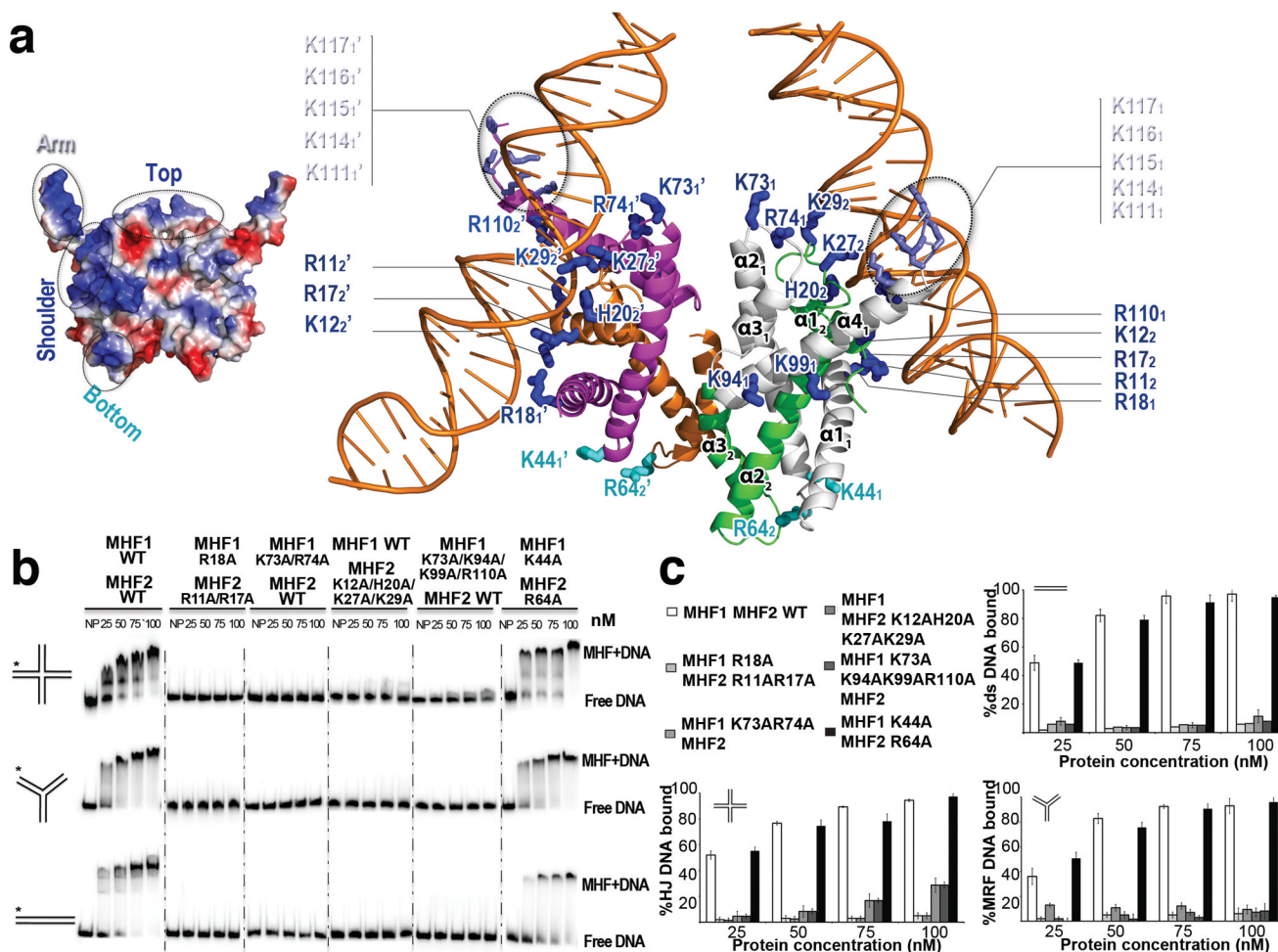


Fig. 2. The positively charged patches of MHF are essential for DNA binding. Overall (top-left inset, shown as electrostatic surface; blue: positive, red: negative) and detailed (ribbon representation) views of the MHF-DNA interface (tetramer 1 in MHF-DNA1). **a**, The critical top and shoulder residues (stick representation) for DNA binding are in blue, the positively charged arm (highlighted by ovals) residues are in light-blue sticks, and those residues that do not affect DNA binding are in cyan. **b**, Electrophoretic mobility shift analysis of MHF mutants with various DNA substrates (the schematics of the HJ DNA, Y-DNA and dsDNA are shown on the left with the radiolabeled terminus indicated by an asterisk). The substrates (30 nM) were incubated with increasing amounts of WT or mutant MHF (25, 50, 75 and 100 nM) and analyzed. NP - no protein added. **c**, The results were quantified and plotted in the bar graphs. Error bars were generated from the standard deviation in triplicate experiments.

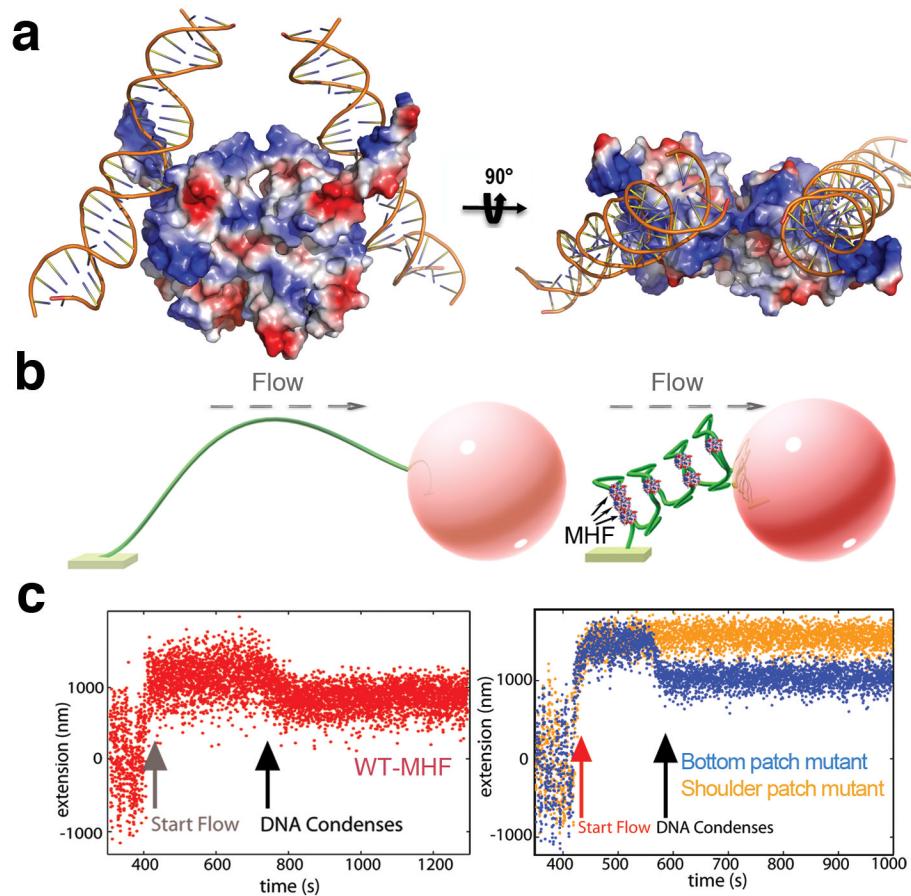


Fig. 3. MHF condenses dsDNA. **a**, Front and top views of the electro-static interaction between the MHF tetramer (Tetramer 1 from MHF-DNA1, surface representation, blue: positive charge, red: negative) and two DNA duplexes. **b**, **c**, Single molecule Tethered Particle Flow Stretching (TPFS) assay of the effect of MHF binding on long dsDNA chains. **b**, Schematic of the TPFS assay of DNA condensation by MHF: the optical beads are shown as pink spheres, the DNA is shown as green strings, MHF tetramers are shown in surface representation, the applied flow is shown in gray arrows. The flow force was calibrated in Supplementary Figure S4. **c**, The dsDNA ligand (4,200 base pairs) was subjected to a fluid flow containing either WT (red), the shoulder patch mutant (R18A/R11A/R17A, blue), or the bottom patch mutant (K44A/R64A, brown) of MHF. The WT and the bottom patch mutant of MHF capable of binding DNA cause a significant shortening of the DNA chain, while the shoulder patch mutant deficient in DNA binding does not.

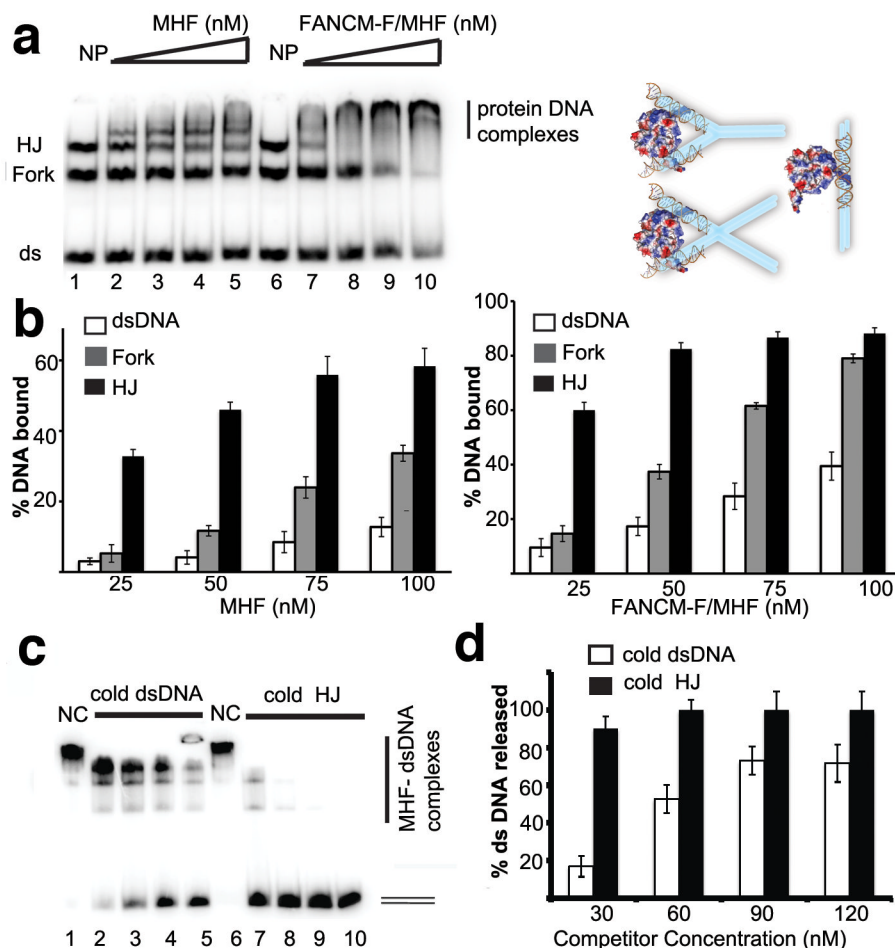


Fig. 4. MHF binds branched DNA preferentially over dsDNA. **a**, Left, analysis of MHF and FANCM-F/MHF complex for binding dsDNA, forked DNA and HJ DNA. Radiolabeled DNA substrates (30 nM each) were incubated with increasing amounts of MHF (Lanes 2–5) or FANCM-F/MHF complex (Lanes 7–10) in the presence of 150 mM KCl. NP - no protein added (Lanes 1 and 6). Right, schematic depicting binding of MHF to forked DNA, HJ, or dsDNA, the charged surface is generated from MHF-DNA1. MHF tetramers are shown as electrostatic surfaces, the observed DNA duplexes are shown as brown helices, and the DNA substrates used in mobility shift assays are illustrated as cyan bars. **b**, Quantification of the results is shown in the bar graphs. Error bars were generated from the standard deviation in triplicate experiments. The left panel shows the data for MHF alone, while the right panel shows the data for the FANCM-F/MHF complex. **c**, Experiments in which a pre-formed MHF-dsDNA nucleoprotein complex was challenged with a cold DNA substrate. MHF complex (100 nM) was preincubated with radiolabeled dsDNA (30 nM). Increasing amounts of unlabeled dsDNA (Lanes 2–5) or HJ (Lanes 7–10) was added (30, 60, 90 and 120 nM). NC - control with no DNA competitor. **d**, Quantification of the results shown in (C). Error bars were generated from the standard deviation in triplicate experiments.

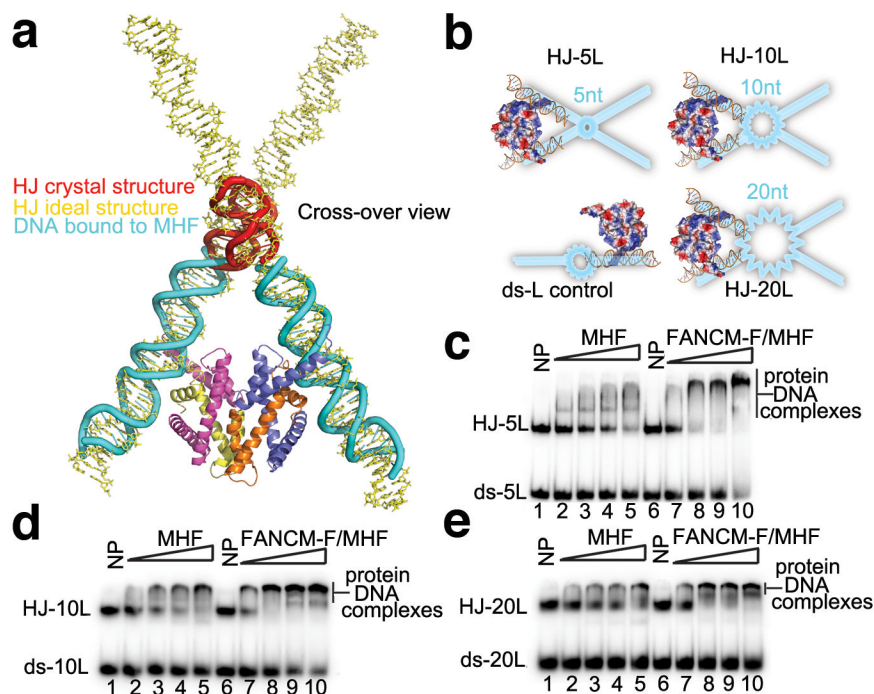


Fig. 5. MHF specifically engages two arms of a DNA junction. **a**, The geometry resemblance between the MHF-bound DNA duplex pair (cyan helices) and the two arms of an ideal Holliday Junction DNA (HJ) (yellow sticks) generated based on the HJ crystal structure (red helices) (PDBID: 3IGT). The DNA-bound MHF tetramer 1 from MHFDNA1 is shown. **b**, Schematic representation of the substrates used. MHF tetramers are shown as electrostatic surfaces, the observed DNA duplexes are shown as brown cartoons, and the DNA substrates are illustrated in cyan with the duplex regions as straight bars and the central bubble regions as jagged lines. **c, d, e**, Analyses of MHF and FANCM-F/MHF for binding the ds-5L and HJ-5L substrates (**c**), the ds-10L and HJ-10L substrates (**d**), and the ds-20L and HJ-20L substrates (**e**). The radiolabeled DNA substrates (30 nM each) were incubated with increasing amounts of MHF (Lanes 2–5) or FANCM-F/MHF complex (Lanes 7–10) in the presence of 150 mM KCl. NP - no protein added (Lanes 1 and 6). Quantification of the results is shown in Figure S4.

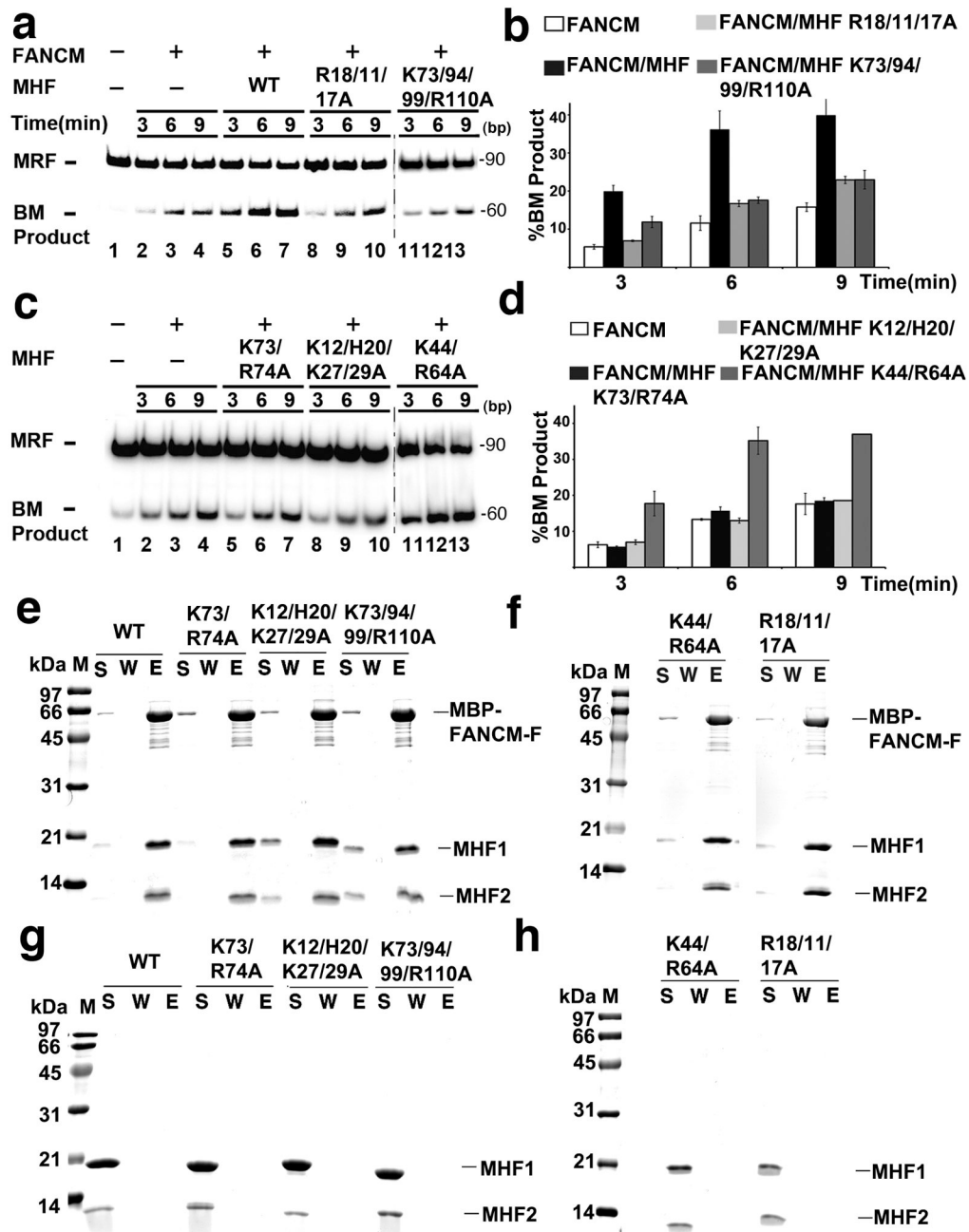


Fig. 6. DNA binding by MHF enhances FANCM-mediated DNA branch migration (BM) but does not correlate with FANCM binding. **a**, FANCM (Lanes 2–4) or FANCM with either wild type (WT) MHF (Lanes 5–7) or MHF mutants (R18/11A/17A) (Lanes 8–10), (K73/94/99/R110A) (Lanes 11–13) was tested for BM activity as a function of time. The concentration of FANCM and MHF (WT or mutant) was 10 nM and of the movable replication fork (MRF) was 5 nM. ATP was present in all the reactions. NP: no protein added. **c**, BM reactions were conducted with FANCM (Lanes 2–4) or with FANCM and MHF mutants (K73/R74A) (Lanes 5–7), (K12/H20/K27/29A) (Lanes 8–10), and (K44/R64A) (Lanes 11–13) as in **a**. **b** and **d**, Quantification of the data shown in **a** and **c**, respectively. Error bars

were generated from the standard deviation in triplicate experiments. **e, f, g** and **h**, Pull-down assays showing that the MHF mutants retain FANCM-binding activity. **e** and **g**, 10 μ g of wild type MHF or MHF mutant as labeled was incubated with the same molar ratio of MBP-His-FANCM-F (FANCM fragment 661-800) before bound to amylose agarose resin (10 μ l; New England Biolabs). 15% of the supernatant (S) and elution (E) fractions, and 2% of total wash (W) fraction were analyzed by SDS-PAGE followed by Coomassie blue staining. **f** and **h**, Control experiments showing that the MHF mutants were not retained on the amylose resin without the inclusion of the MBP-tagged FANCM-F.

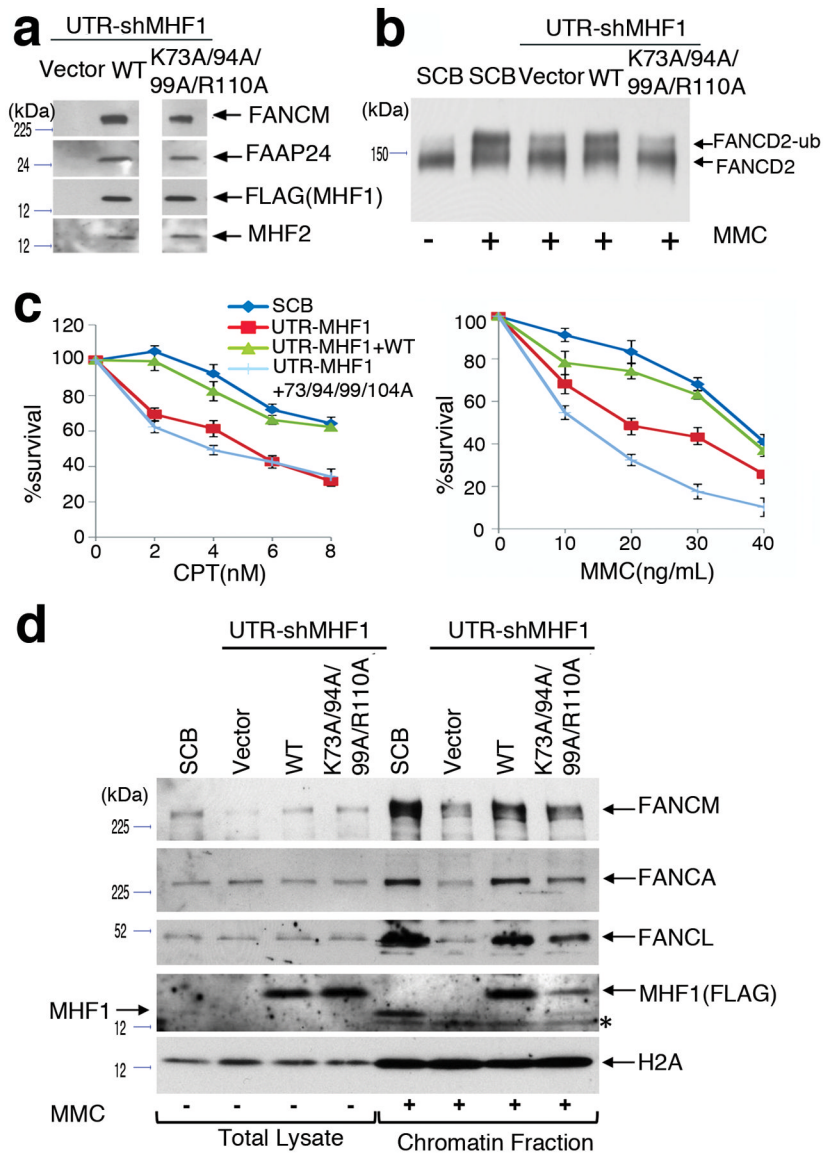


Fig. 7. Cellular phenotypes of MHF DNA-binding mutants. **a**, The MHF1 DNA-binding mutant (K73A/K94A/K99A/R110A) is proficient in interactions with MHF2 and FANCM-FAAP24. HeLa cells transduced with lentiviruses carrying MHF1-UTR were further transduced with retroviruses carrying empty vector (Lane 1), Flag-WT-MHF1 (Lane 2) or Flag-MHF1 mutant (Lane 3), and the anti-Flag immuno-precipitated protein complex was analyzed by immunoblotting with the indicated antibodies. **b**, The MHF1 DNA-binding mutation (K73A/K94A/K99A/R110A) affects FANCD2 monoubiquitination. Immunoblot showing impaired FANCD2 monoubiquitination in the MHF1 mutant. HeLa cells were transduced with lentiviruses carrying either control shRNA (SCB, Lanes 1–2) or MHF1-UTR (UTR-shMHF1, Lanes 3–5). Then, HeLa-MHF1-UTR cells were transduced with retroviruses carrying either empty vector (vector, Lane 3), shRNA resistant WT-MHF1 cDNA (WT, Lane 4), or mutant-MHF (K73A/K94A/K99A/R110A, lane 5). Cells were treated with 100ng/ml MMC for 16h, and blotted by anti-FANCD2 antibody. FANCD2-ub: monoubiquitinated FANCD2. **c**, DNA binding by MHF is required for cellular resistance to

DNA damaging agents. HeLa cells with endogenous MHF1 depleted by shRNA treatment and expressing either wild type or the DNA-binding mutant of MHF1 (see **b** for details) were treated with the indicated concentration of CPT or MMC. Visible colonies from 200 cells were counted after 10 days. The data represent the percent survival, as compared with untreated cells. Each experiment was performed in triplicate, and mean values are shown with standard deviations. **d**, Immunoblot showing the levels of FA core complex members in total lysate and chromatin fraction. HeLa cells were transduced with lentiviruses carrying either control shRNA (SCB) or MHF1-UTR (UTR-shMHF1). Then, HeLa-MHF1-UTR cells were transduced with retroviruses carrying empty vector (vector), shRNA resistant WT-MHF1 cDNA (WT) or mutant-MHF (K73A/K94A/K99A/R110A). Cells were treated with 100ng/ml MMC for 16h, followed by cellular fractionation into total lysate and chromatin fraction. The relevant proteins were revealed by immunoblotting. H2A serves as loading control and a marker for the chromatin fraction. Asterisk (*) denotes a nonspecific cross-reactive band.

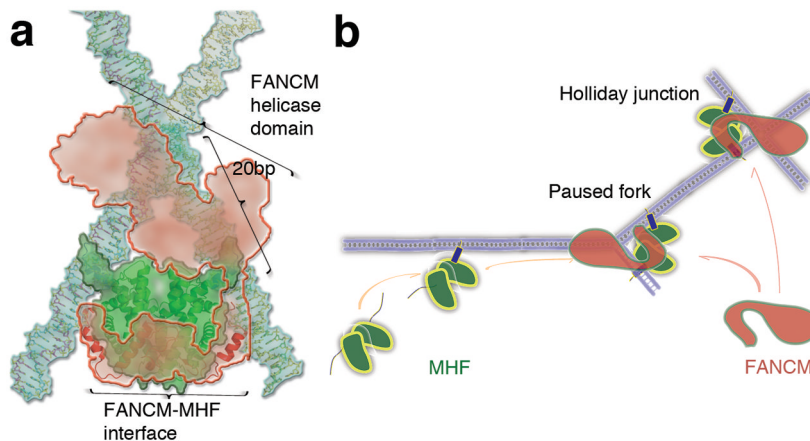


Fig. 8. Model explaining the function of MHF in the FA pathway of DNA damage repair. **a**, The MHF-FANCM (661-800) structure (PDB ID: 4DRB) is overlaid onto the tetramer 1 in the MHF-DNA complex (FANCM in red, MHF in green, and DNA in cyan). The ideal Holliday junction DNA is shown as sticks. The homology-modeled helicase domain of FANCM is shown as semi-transparent red contour. **b**, Model depicting a role of MHF (green) in facilitating FANCM (red) to branched DNA (purple) in FA pathway of DNA repair. The arm regions are also displayed before (as blue wires) or after (or blue tailed-tubes) DNA binding.

Table 1

Data collection and refinement statistics

	MHF-apo1	MHF-apo2	MHF-DNA2	MHF-DNA1	MHF-Se
Data collection					
Space group	P2₁	C2	P3	P3	P1
Cell dimensions					
<i>a, b, c</i> (Å)	41.08 128.76 88.84 90.00 100.96 90.00	61.97 52.94 60.28 90.00 114.91 90.00	252.39 252.39 131.36 90.00 90.00 120.00	250.13 250.13 65.58 90.00 90.00 120.00	40.73 40.82 59.48 91.90 105.55 95.47
<i>a, b, g</i> (°)					
Wavelength(Å)	1	1	0.9792	0.9792	0.9795
Resolution (Å)	50-2.5 (2.6-2.5)	50-1.8 (1.9-1.8)	50-6.5 (6.7-6.5)	50-7.2 (7.4-7.2)	37-2.1 (2.2-2.1)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.119 (***#)	0.046 (***#)	0.130 (***#)	0.131 (***#)	0.073 (0.71)
<i>I</i> / <i>σ</i>	17.5 (1.9)	29.1 (1.5)	6.0 (0.5)	8.0 (1.0)	8.4 (1.1)
Completeness (%)	99.9 (100.0)	99.8 (99.9)	99.9 (99.9)	99.8 (99.7)	91.7 (76.1)
Redundancy	11.4	5.1	2.7	3.1	3.5
<i>Wilson B</i>	35	29	388#	393#	35
<i>Twin_law</i> (winning fraction %)	-	-	<i>k,h,-l</i> (0.49)	<i>k,h,-l</i> (0.44)	-
Refinement					
Resolution (Å)	50-2.5	50-1.8	50-6.5	50-7.2	37-2.1
No. reflections	31307	16358	18426	6243	20776
<i>R</i> _{work} / <i>R</i> _{free}	0.223/0.264	0.231/0.25	0.284/0.313	0.268/0.288	0.230/0.262
No. atoms					
Protein/DNA	5332	1327	31827	15914	2666
Water	147	56	N/A	N/A	2745
<i>B</i> -factors					
Protein/DNA	45.4	43.4	316	281	56.8
Water	39.2	42.8	N/A	N/A	47.9

	MHF- <i>apo</i> 1	MHF- <i>apo</i> 2	MHF-DNA2	MHF-DNA1	MHF-Se
R.m.s. deviations					
Bond lengths (Å)	0.008	0.007	0.010	0.007	0.009
Bond angles (°)	1.0	0.9	1.5	1.1	1.0
Ramachandran plot					
Most favored (%)	100	100	98	99	100
Allowed (%)	0	0	1.97	0.94	0
Outliers (%)	0	0	0.03	0.06	0

* Values in parentheses are for highest-resolution shell

---# Values are above 1.

♣ Wilson statistics at low resolution is not accurate.

♠ Twinning fractions estimated from the twin refinement in Refmac5 37