Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwigs-Maximilians Universität München

Structural and Biochemical Characterization of Human GEN1 Holliday Junction Resolvase

Shun-Hsiao Lee
aus
Hong Kong

2017

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Frau Prof. Dr. Elena Conti betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 5.12.2017

.....

Shun-Hsiao Lee

Dissertation eingereicht am 22.09.2017

Erstgutachter: Prof. Dr. Elena Conti

Zweitgutachter: Prof. Dr. Karl-Peter Hopfner

Mündliche Prüfung am 21.11.2017

Summary

Homologous recombination (HR) is a universal mechanism found in all domains of life for DNA segregation as well as for "error-free" DNA double-strand break (DSB) repair. The key intermediate in this process is a DNA four-way junction also known as a Holliday Junction (HJ), which is formed after pairing of homologous DNA and strand exchange. In later stages of HR, joint DNA molecules have to be faithfully eliminated in order to avoid improper chromosome segregation and to maintain the genome integrity. Human GEN1 is a member of the Rad2/XPG nuclease family and it is the classical Holliday junction-resolving enzyme. Interestingly, this enzyme has no sequence homology to the well-characterized resolvases found in prokaryotes but shares the same function in symmetric HJ cleavage. In contrast to other members of the Rad2/XPG family, GEN1 can form a homodimer and recognize DNA four-way junctions as well as 5′ flaps, replication forks and splayed arms. Its roles in removing recombination and replication intermediates have been implicated *in vivo*. However, many functional mechanisms about this eukaryotic enzyme are still elusive. This thesis presents a structural and biochemical characterization of human GEN1. The X-ray structure of the GEN1-DNA complex revealed that the enzyme contains a chromodomain C-terminal of the conserved Rad2/XPG nuclease core. The chromodomain has direct interactions with DNA and is critical for efficient substrate recognition and cleavage. Further biochemical studies defined the structural elements for substrates discrimination. In brief, human GEN1 has a helical arch that fosters the cleavage of 5′ flaps but has minor effects on HJs, suggesting a conserved catalytic mechanism resembling the one of flap endonuclease from bacteriophage T5. In addition, an unstructured positively charged cluster was identified C-terminal of the chromodomain, which is required for proper substrate cleavage at near physiological salt concentrations. Lastly, biophysical analyses of GEN1-substrate interactions confirmed that GEN1 works as a monomer on 5′ flaps and dimerizes upon binding to HJs. A model is proposed that GEN1 first targets DNA via its positively charged cluster, and adapts in a flexible way to discriminate different DNA structures. For HJs, GEN1 forms homodimer and coordinates the symmetrical dual incisions. As a 5′ flap endonuclease, GEN1 remains in monomeric form and the dimerization-triggered catalysis is bypassed.

Preface

The work in thesis was performed in the laboratory of Dr. Christian Biertümpfel at the Max Planck Institute of Biochemistry, Martinsried, Germany. This thesis comprises two main topics on human Holliday junction resolvase GEN1 and led to one first-author publication and a first-author full manuscript ready for submission. Therefore, this thesis is presented in the cumulative style. Chapter 1 gives a general introduction to the whole field and briefly summarize the current research. Chapter 2 presents the published structure of the GEN1-DNA complex and the features of GEN1's chromodomain. Chapter 3 focuses on the biochemical characterization of GEN1's substrate recognition. Finally, Chapter 4 gives a general discussion on both topics, the relevance of the findings and future directions for the project.

TABLE OF CONTENTS

LIST OF FIGURES

Introduction to Eukaryotic Joint Molecule Resolution

Structure of Human GEN1-DNA Complex: The Chromodomain is required for Efficient Substrate Recognition and Cleavage

Two-mode DNA Recognition by Human Holliday Junction Resolvase GEN1

Concluding Remarks

LIST OF TABLES

Chapter 1 –
Introduction to Eukaryotic Joint Molecule Resolution

1.1. Homologous recombination

Homologous recombination (HR) is a conserved mechanism in all three domains of life. It is a powerful tool to repair DNA double-stranded breaks (DSBs) arising from ionizing radiation, metabolic reactive oxygen species (ROS) and the processing of inter-strand crosslinks (ICLs) and broken replication forks. The essence of this system is to use homologous sequences as the template to restore the genetic information in an error-free fashion **(Figure 1-1)** (Heyer, 2015; Jasin and Rothstein, 2013; Kowalczykowski, 2015; San Filippo et al., 2008). In general, cells prefer using a sister chromatid as the template for homologous recombination in order to avoid the loss of heterozygosity (LOH). Therefore, HR is only activated in late S phase and G_2 phase when newly replicated chromosomes are available, the so-called "mitotic recombination" (Ira et al., 2004). On the other hand, in meiotic cells, specialized "meiotic recombination" happens between homologous chromosomes. Programmed DNA double-stranded breaks are induced by SPO11 and its associated proteins to initiate the recombination (Bergerat et al., 1997; Keeney et al., 1997). The strand exchange between homologs potentially introduces gene diversity to the offspring and the crossovers are generated to facilitate chromosome segregation during meiosis I. DNA double-stranded breaks that arise during G_1 phase are primarily repaired by the alternative pathway named non-homologous end-joining (NHEJ). The pathway choice between HR and NHEJ is tightly regulated by the actions involving the proteins BRCA1 and 53BP1 in a cell cycle-dependent manner (Daley and Sung, 2014; Symington and Gautier, 2011).

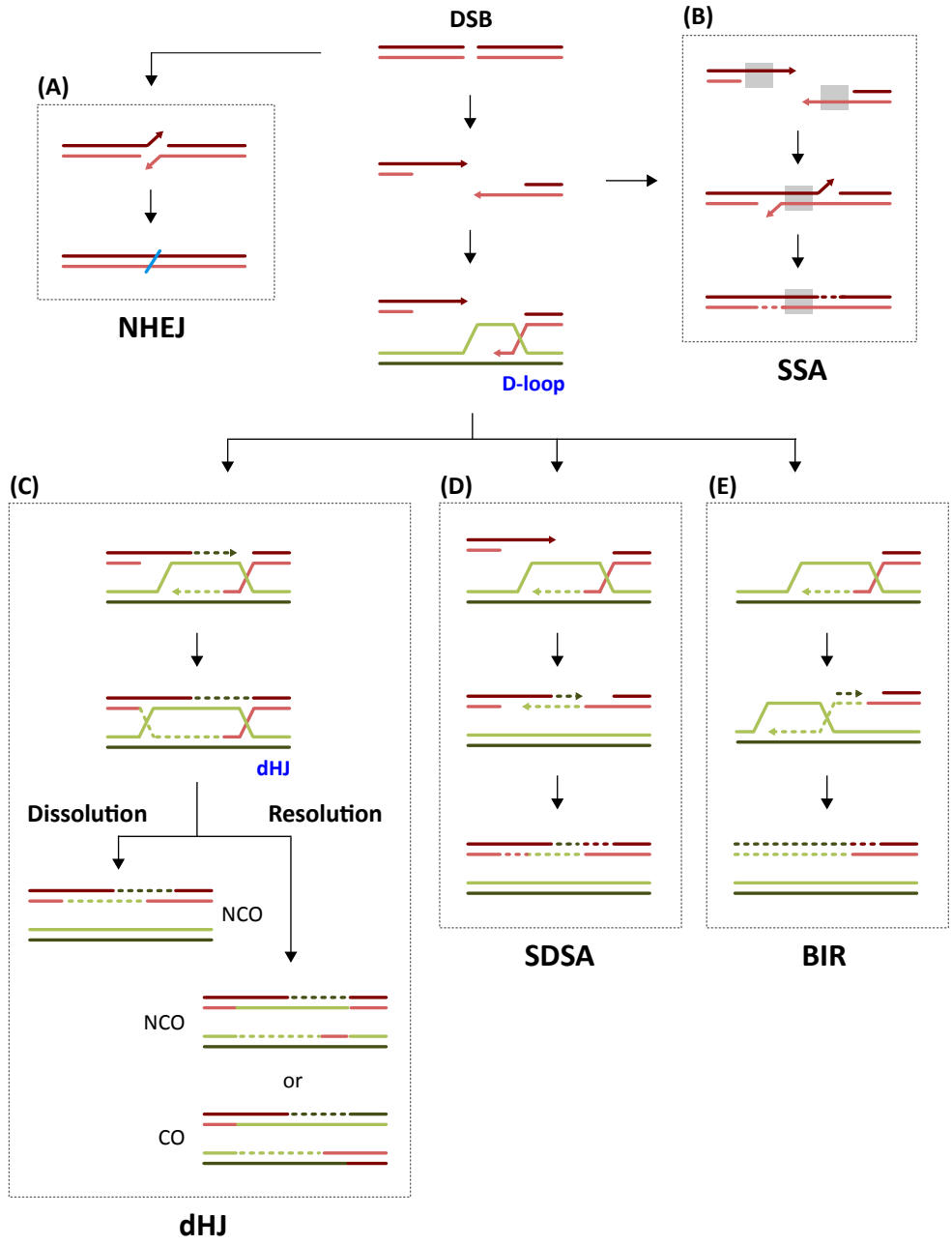


Figure 1-1. Pathways of eukaryotic DNA double-stranded break (DSB) repair.

(A) Non-homologous end joining (NHEJ) repairs broken DNA through the direct ligation after end-processing. (B) Single-stranded annealing (SSA) involves the pairing of flanking homologous repeats, resulting in deletion of the intervening DNA sequences. (C) The double Holliday junction (dHJ) pathway involves a second-end capture after D-loop formation. Dissolution of dHJs yields exclusively non-crossover (NCO) outcomes and the resolution of HJs generates both crossover (CO) and NCO products. (D) In synthesis-dependent strand annealing (SDSA), DNA synthesis on the D-loop heteroduplex eventually leads to its dissociation and the newly synthesized strand re-anneals with the original duplex. (E) Break-induced replication (BIR) repairs the broken replication forks. DNA synthesis restarts by using the homologous strand as template, causing loss of heterozygosity. The figure is adapted from (Mehta and Haber, 2014).

1.1.1. End resection

To initiate homologous recombination, the DNA ends must be nucleolytically resected to form single-stranded 3′ overhangs (Blackwood et al., 2013; Mimitou and Symington, 2011). The long 3′ single-stranded DNA (normally thousands of bases) serves as a platform for subsequent loading of regulators and ensures that the pathway does not fall back to NHEJ. In eukaryotes, the MRN/X complex (MRE11-RAD50-NBS1 in humans, Mre11-Rad50-Xrs2 in budding yeast) together with the associated factor carboxyl terminal binding protein interacting protein (CtIP, Sae2 in budding yeast) is the key machinery to process the DNA ends (Symington and Gautier, 2011). It has been shown that the MRN/X complex is needed for removing the SPO11- and topoisomerase II-mediated DNA adducts (Furuse et al., 1998; Hartsuiker et al., 2009). MRE11 is the nuclease subunit that is in charge of the DNA digestion. Rad50 is an ABC-type ATPase that promotes Mre11 activity and provides a DNA-tethering function. Mre11-Rad50 together is the core complex that is conserved from prokaryotes to eukaryotes (Blackwood et al., 2013). NBS1/Xrs2 functions as a regulatory module that interacts with associated proteins and enhances the activity of MRE11 and RAD50 in eukaryotes (Lee et al., 2003). The current model suggests that phosphorylated CtIP/Sae2 promotes the cryptic endonuclease activity of MRE11 to introduce a nick proximal to the protected DNA ends, and then the MRN/X complex nucleolytically degrades single-stranded DNA using its 3′–5′ exonuclease activity (Anand et al., 2016; Cannavo and Cejka, 2014). The further extensive processing is achieved by EXO1 (exonuclease 1) and DNA2-BLM/Sgs1 (DNA replication ATP-dependent helicase/nuclease 2, Bloom's syndrome helicase in humans, slow growth suppressor 1 in budding yeast) in 5′–3′ direction (Cannavo et al., 2013; Cejka et al., 2010a; Mimitou and Symington, 2008; Nimonkar et al., 2011; Nimonkar et al., 2008; Niu et al., 2010; Zhu et al., 2008).

1.1.2. Strand invasion and D-loop formation

Once the DNA ends are processed, the exposed 3′ single-stranded tail is coated with replication protein A (RPA), which is a heterotrimeric protein complex that specifically binds to ssDNA with high affinity. The binding of RPA can remove DNA secondary structures and protects ssDNA from unwanted degradation. RPA is then replaced by the Rad51 recombinase, forming a "nucleoprotein filament" or "presynaptic filament" that is the key element for homology search **(Figure 1-2)**. Rad51 is the homolog of bacterial recombinase

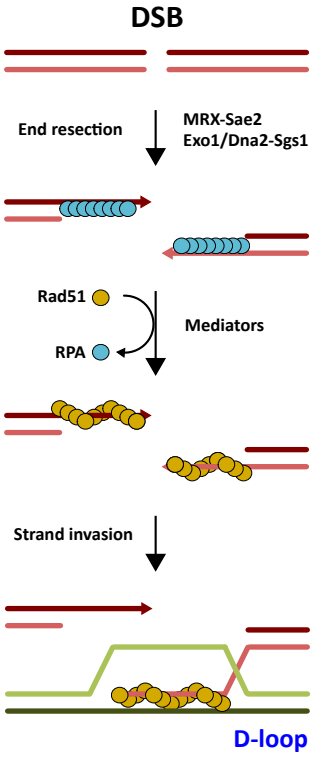


Figure 1-2. Strand invasion and D-loop formation.

To initiate HR, DNA ends have to be extensively resected to form the 3' overhang tails. Exposed ssDNA is first protected by replication protein A (RPA), and then replaced by Rad51 recombinase via mediators. The Rad51 coated "presynaptic filament" can invade homologous sequences and form the displacement loop, or D-loop. Figure is adapted and redrawn from (Mimitou and Symington, 2009).

RecA that comprises ATPase activity (Aboussekhra et al., 1992; Basile et al., 1992; Shinohara et al., 1992). Biochemical studies have shown evidence that ATP binding is essential for the formation of the presynaptic filament, and ATP hydrolysis is involved in Rad51 dissociation and turnover. Human RAD51 forms a heptameric complex in solution, when it assembles on DNA, it creates a right-handed helical structure that can extend to thousands of bases. The DNA bound with Rad51 is stretched about 150% to the B-form conformation (Ogawa et al., 1993). The replacement of RPA by Rad51 is a very slow process due to the fact that RPA has a stronger affinity to ssDNA. Thus, this reaction requires certain cofactors for enhancement. These proteins involved in facilitating Rad51 coating are termed recombination mediators. Rad52 is the predominant mediator in *S. cerevisiae* (Sung, 1997). It forms a seven-subunits ring in solution and has direct interactions with Rad51 and RPA that recruits the recombinase to the RPA-coated ssDNA and has a nucleation effect to the presynaptic filament formation. However, in human cells, RAD52 has weak mediator activities. Instead, another protein, breast cancer type 2 susceptibility protein (BRCA2), plays a critical role (Jensen et al., 2010; San Filippo et al., 2006; Yang et al., 2005). Similar to RAD52, BRCA2 interacts with RPA and loads RAD51 onto ssDNA. An additional function of BRCA2

promoting RAD51 binding to ssDNA instead of dsDNA has also been documented (Yang et al., 2002).

The presynaptic filament is the only structure that can lead to homology pairing (Sung and Robberson, 1995). From studies on the prokaryotic RecA protein it is clear that the presynaptic filament first randomly searches for dsDNA (Bianco et al., 1998). Once homology is found, the ssDNA pairs with the homologous sequence and forms a paranemic, three-stranded intermediate. This Rad51-ssDNA-dsDNA structure is called "synaptic complex", and strand extension is further stabilized and creates the plectonemic joint, "displacement loop" or "D-loop" (Chi et al., 2007; Pezza et al., 2007). From now on, DNA polymerases can utilize the 3'-hydroxyl group as the primer and extend the DNA by using the donor strand as template.

The D-loop is the key decision point for the subsequent pathway choice. Several models have been suggested to use the D-loop structure as a common precursor. These pathways are double Holliday junction (dHJ), synthesis-dependent strand annealing (SDSA) and break-induced replication (BIR) (Kowalczykowski, 2015; Mehta and Haber, 2014). Brief introductions to each pathway are summarized below.

1.1.3. Double Holliday junction (dHJ) mechanism

In 1964, Robin Holliday proposed a model to describe gene conversion during meiosis (Holliday, 1964). This model suggests that after replication, DNA single-stranded nicks allow strand exchange across homologous sequences, creating a cross-stranded four-way junction. The genetic information can be transferred to another homolog via the migration of the branch point. Even though this model has been modified and corrected (now we know that recombination is initiated with DSBs and that four-way junctions adopt an anti-parallel conformation), the central idea is still valid after the extensive studies for more than half century. The four-way junction structures, or "Holliday junctions (HJs)", have been observed not only in meiotic cells but also in mitotic cells, suggesting that it is one of the key intermediates in genetic recombination (Bzymek et al., 2010; Cromie et al., 2006; Schwacha and Kleckner, 1995).

The modified "double Holliday junction (dHJ)" mechanism is one of the widely accepted models for DNA double-stranded break repair (Szostak et al., 1983). The formation of dHJs requires another end capture after the D-loop intermediate is matured following heteroduplex extension **(Figure 1-1, panel C)**. This can be achieved by annealing the second 3′ single-stranded overhang from the other side of the DSB to the homologous DNA that is replaced by the first strand invasion. DNA synthesis from both 3′-OH primers fills up the gap and the double Holliday junction is eventually created. This structure can be "dissolved" by a topoisomerase-helicase driven reaction to generate non-crossover (NCO) products or "resolved" by structure-selective endonucleases to yield both crossover (CO) and non-crossover outcomes. The details of joint-molecule disentanglement will be discussed in later sections.

1.1.4. Synthesis-dependent strand annealing (SDSA)

Alternatively, a D-loop structure can go through another pathway termed synthesis-dependent strand annealing **(Figure 1-1, panel D)**. This model was first proposed to explain the contradiction of the dHJ model that the predicted ratio of crossover/non-crossover configurations did not match to the observation (Allers and Lichten, 2001; Ferguson and Holloman, 1996; Nassif et al., 1994; Resnick, 1976). In the SDSA model, the second 3′ overhang does not pair to the homologous chromosome. Instead, the matured D-loop is dissociated and the extended invasion strand anneals with the 3′ overhang from the original duplex. Since this mechanism does not involve HJ formation and the outcomes are exclusively non-crossovers, it is suggested to be the key pathway for mitotic cells. Genetic studies have identified that several anti-recombinational helicases participate in the pathway choice between dHJ and SDSA, such as Srs2, Sgs1 and Mph1 in *S. cerevisiae* and RTEL1 in *C. elegans* (Barber et al., 2008; Ira et al., 2003). Presumably these anti-recombination helicases disassemble D-loops via branch migration and therefore promote the non-crossover pathway.

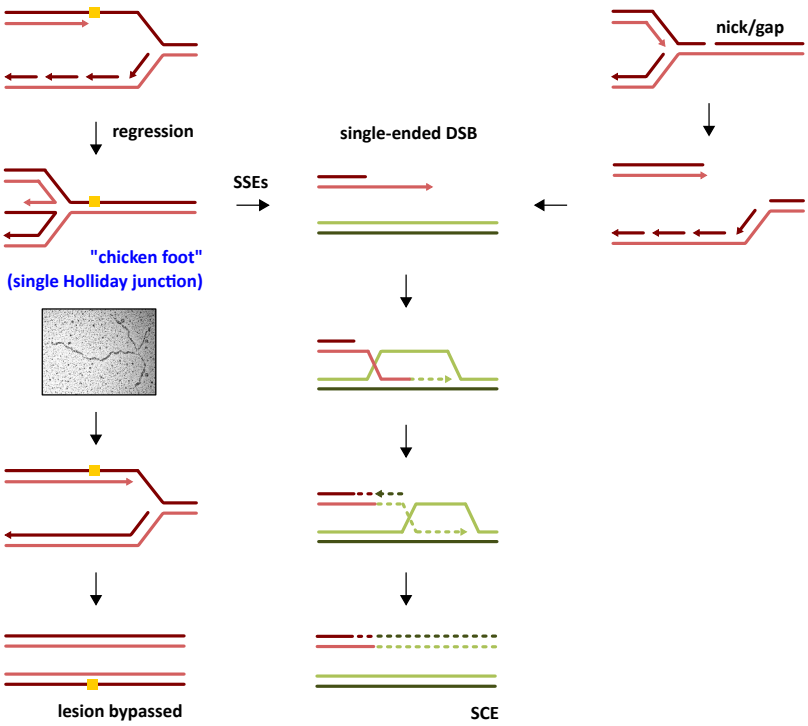


Figure 1-3. Repair of stalled and collapse replication forks.

The reversion of stalled replication forks generates chicken foot structures, or single Holliday junctions (sHJs). Replication can restart by re-establishing the fork structure, and the lesion is bypassed. On the other hand, processing of sHJs by resolvases leads to single-ended DSBs, which can be further repaired by break-induced replication (BIR). Moreover, unrepaired nicks or gaps during replication also cause single-ended DSBs. BIR uses a sister chromatid to restore the genetic information. Therefore, the level of sister-chromatid exchanges (SCEs) is elevated. The figure is adapted from (Mehta and Haber, 2014). The image of a HJ is taken from (Zellweger et al., 2015).

1.1.5. Break-induced replication (BIR)

Certain types of DNA lesions could be developed as double-stranded breaks with only one DNA end. When single-stranded breaks remain unrepaired during S phase, a replication fork will collapse at nick or gap sites and generate a broken duplex **(Figure 1-3)** (Kuzminov, 2001).

In such circumstances, the broken end can be processed and invade to the other intact DNA using the same mechanism as D-loop formation. This restart mechanism for collapsed replication forks is termed “break-induced replication” (Malkova et al., 1996). In other situations, replication could stop by roadblocks such as pyrimidine dimers. The stalled

replication forks can go through a regression mechanism that remodels the fork structure (Postow et al., 2001). The stalled strand recaptures the other newly synthesized strand and uses it for extension, forming a single Holliday junction or "chicken foot" structure. The damage is bypassed and the replication restarts when the two synthesized strands re-anneal to the original templates. On the other hand, the single Holliday junction can be a substrate for structure-selective endonucleases (SSEs), such as Mus81-Mms4 and Yen1 in yeast (Tay and Wu, 2010). The cleavage products can re-initiate the replication by BIR pathway (Hanada et al., 2007; Petermann and Helleday, 2010).

1.2. Joint molecule disentanglement in eukaryotes

The joint molecules, especially Holliday junctions, arise during homologous recombination are important for gene conversion. However, these physical linkages are toxic for chromosome segregation and therefore have to be faithfully removed to maintain genome integrity. In prokaryotes, cells encode a group of endonucleases, called "resolvases", which specifically target to joint molecules and resolve branched structures. Eukaryotes preserve the resolution system but developed a much more sophisticated regulation mechanism (Wyatt and West, 2014). On top of that, another "dissolution" system has been identified (Bizard and Hickson, 2014). These systems are well controlled and coordinated in timely and spatially manners to safeguard genome stability. This section will summarize the regulation and mechanisms of both systems in general, and a detailed introduction for the main character of this thesis, GEN1, will be emphasized in the section 1.3.

1.2.1. Double Holliday junction dissolution

In the classical Holliday junction model, four-way junctions are resolved by structure-selective endonucleases. Depending on the cleavage orientation, both non-crossovers (NCOs) and crossovers (COs) are the possible outcomes. However, somatic cells prefer to avoid CO as it potentially leads to sister chromatid exchanges (SCEs) and loss of heterozygosity (LOH). Eukaryotes evolved an alternative "dissolution" system to reduce unwanted COs. This reaction is driven by the "dissolvasome" or BTR/STR complex (BLM helicase-TOP3α-RMI1-

RMI2 in humans, Sgs1-Top3-Rmi1 in budding yeast) with the combined actions of a helicase and a topoisomerase **(Figure 1-4)**.

Genetic studies in *S. cerevisiae* have revealed that the RecQ family 3'-5' helicase Sgs1 (slow growth suppressor 1) and the type IA topoisomerase Top3 (topoisomerase III) are involved in recombination and play a role in suppressing CO formation (Gangloff et al., 1994; Ira et al., 2003; Wallis et al., 1989; Yamagata et al., 1998). Similar functions have been observed for orthologs in higher eukaryotes (Chaganti et al., 1974; Li and Wang, 1998; Seki et al., 2006). The third component, Rmi1 (RecQ-mediated genome instability protein 1, RMI1 in humans), was identified by its genetic association with Sgs1 and Top3. It has been shown that these three proteins associate with each other independently to form a stable complex (Chang et al., 2005; Mullen et al., 2005). Biochemical studies demonstrated that purified human BLM helicase (Bloom's syndrome helicase, the ortholog of budding yeast Sgs1) together with

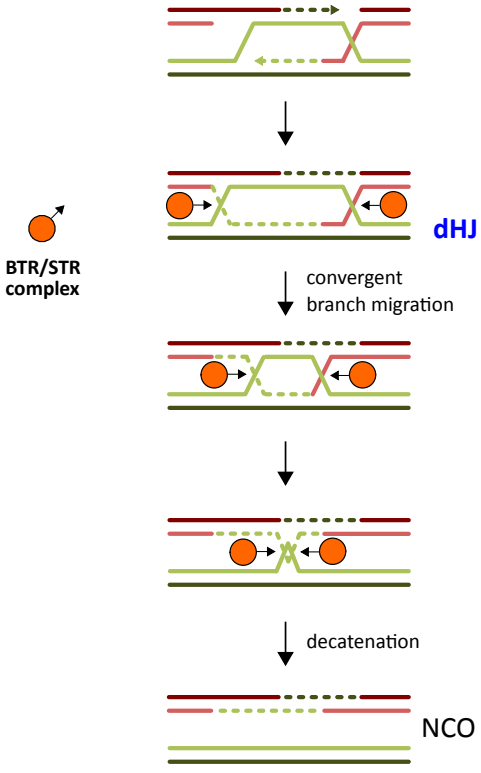


Figure 1-4. Double Holliday junction dissolution.

Double Holliday junctions (dHJs) can be dissolved by BTR/STR complex (BLM-TOP3α-RMI1-RMI2 in humans, Sgs1-Top3-Rmi1 in budding yeast). The reaction is carried out by the combined action of a 3'-5' helicase and a type IA topoisomerase. BTR/STR complex promotes the convergent branch migration of dHJs and generates non-crossover (NCO) products.

TOP3α (topoisomerase IIIα, the ortholog of yeast Top3) are capable of dissolving dHJ substrates (Wu and Hickson, 2003). The disentanglement of dHJs is driven by a convergent branch migration mechanism, in which the protein machinery pushes two junctions migrating inward and eventually collapsing them and generating exclusively NCO products (Cejka et al., 2010b; Plank et al., 2006). The reaction requires the helicase activity of BLM/Sgs1 to promote branch migration, and TOP3α/Top3 releases the topological stress and decatenates the final intermediate through its single-stranded DNA passage activity. RMI1/Rmi1 functions as a regulatory protein that has no enzymatic activity. It has been shown that Rmi1 dramatically improves the dissolution efficiency (Cejka et al., 2010b). However, the regulation mechanism of Rmi1 is unclear. Biochemical and structural studies proposed that Rmi1 promotes dissolution by regulating the catalytic dynamics of Top3 (Bocquet et al., 2014; Cejka et al., 2012). Another associating factor, RMI2, which only exists in humans has an essential role in dissolution *in vivo*, but the mechanism is still poorly understood (Singh et al., 2008; Xu et al., 2008).

1.2.2. Holliday junction resolution

Holliday junction resolution has been extensively studied in bacteria, archaea and bacteriophages. The endonucleases recognizing and cleaving Holliday junctions are called "resolvases", such as *E. coli* RuvC, T4 endonuclease VII and T7 endonuclease I. Despite the sequence and structural diversity among these enzymes, common features have been preserved: (1) the proteins assemble stable homodimers in solution, (2) they bind with high affinity to Holliday junctions *in vitro*, and (3) they introduce symmetrical incisions into four-way junctions, resulting in two nicked duplexes which can be readily resealed by ligases (Bennett et al., 1993) **(Figure 1-5, panel A)**. Depending on the cleavage orientation, processing of dHJs generates both CO and NCO products **(Figure 1-5, panel B)**.

Eukaryotes evolved a much more sophisticated system for HJ resolution. Several enzymes have been implicated as resolvases, namely GEN1, MUS81-EME1 and SLX1-SLX4 in humans. GEN1 is the only enzyme, which is considered as a canonical resolvase as its biochemical properties are similar to the prokaryotic prototypes. MUS81-EME1 and SLX1-SLX4, however, are the "non-canonical resolvases" since their cleavage mechanisms are apparently distinct from the well-studied prokaryotic resolvases. This section will focus on the non-canonical resolvases, and the classical resolvase GEN1 will be highlighted in the section 1.3.

1.2.2.1. MUS81-EME1 complex

MUS81 (MMS and UV sensitive protein 81) and its associated partner EME1 (essential meiotic endonuclease 1) belong to the ERCC4/XPF (excision repair complementing defective in chinese hamster/Xeroderma pigmentosum group F-complementing protein) nuclease family comprising a conserved ERCC4 nuclease domain. While only MUS81 preserves the essential catalytic GDX_nERKX_3D motif, EME1 is catalytic dead (Ciccia et al., 2008). Mus81-Eme1 has been originally proposed as a HJ resolvase based on its essential meiotic role in *S. pombe* (Boddy et al., 2001). Chromosome mis-segregation and genome instability were observed in cells lacking *mus81*. These meiotic phenotypes can be rescued by the ectopic expression of the bacterial resolvase RusA, supporting the proposed function of Mus81-Eme1 in HJ processing (Boddy et al., 2001). Biochemical studies further confirmed that human MUS81-EME1 is capable of resolving model HJ substrates (Chen et al., 2001; Ciccia et al., 2003). However, the cleavage products are rather distinct from the ones of canonical resolvases. MUS81-EME1 makes two asymmetric incisions near the junction point generating two linear DNA products with one carrying a flap and the other bearing a gap. These products cannot be direct substrates for ligases therefore further processing is needed (Constantinou et al., 2002). In fact, the enzyme efficiency to HJs is relatively low. Instead, MUS81-EME1 is a robust enzyme toward replication forks, 3′ flaps, nicked Holliday

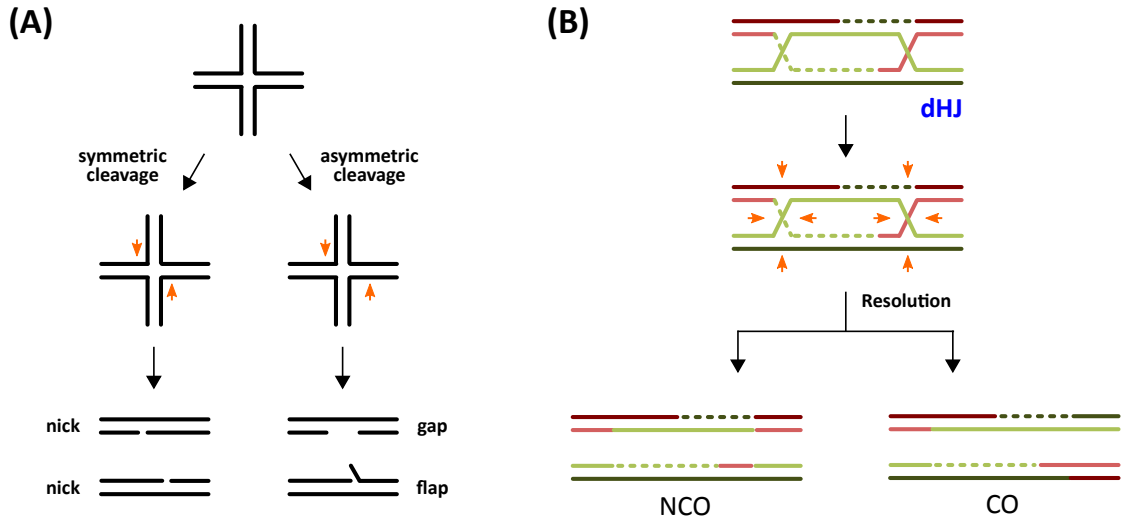


Figure 1-5. Holiday junction (HJ) resolution.

(A) HJs can be processed by junction-resolving enzymes. Coordinated symmetric incisions generate two nicked duplexes that can be directly ligated. Asymmetric cuts on HJs create gapped and flap products, which have to be further processed. (B) Double Holliday junction resolution generates both crossover (CO) and non-crossover (NCO) outcomes. Nucleolytic incisions are presented as orange arrows.

junctions (nHJs) and D-loop structures, leading to the hypothesis that MUS81-EME1 targets to other DNA structures arising during recombination *in vivo* (Ciccia et al., 2003; Hollingsworth and Brill, 2004; Osman et al., 2003; Whitby et al., 2003).

1.2.2.2. SLX1-SLX4 complex

SLX1 and SLX4 were identified in *S. cerevisiae* by a synthetic lethal screening while searching for the genes that are essential for the cells lack of the Sgs1 helicase (SLX refers to synthetic lethal of unknown function) (Mullen et al., 2001). Sequence analysis revealed that SLX1 is a nuclease belonging to the GIY-YIG family, which is related to the bacterial UvrC protein. SLX4 is a scaffold protein that comprises multiple domains and orchestrates many biological functions by interacting with different partners. Besides binding to SLX1, it also serves as a docking platform for MUS81-EME1, XPF-ERCC1, MSH2-MSH3, and TRF2-TRF2IP, which plays an important role in recombination, repair of inter-strand crosslinks (ICLs), mismatch repair (MMR) and telomere maintenance, respectively (Andersen et al., 2009; Fekairi et al., 2009; Munoz et al., 2009; Svendsen et al., 2009). The nuclease activity of SLX1 requires the binding of SLX4, cleavage of HJs, 5' flaps, 3' flaps, replication forks and stem loop structures have been observed (Fekairi et al., 2009; Fricke and Brill, 2003; Munoz et al., 2009; Svendsen et al., 2009). Even though SLX1-SLX4 is capable of generating symmetric incisions into HJs, the cuts are uncoordinated hence the specificity is low and only minor parts of the products can be directly ligated (Svendsen et al., 2009; Wyatt et al., 2013). Therefore, it is still controversial to classify SLX1-SLX4 as a canonical HJ resolvases. In fact, more recent studies indicated that SLX1-SLX4 processes HJs in an alternative mechanism by cooperating with MUS81-EME1 (Castor et al., 2013; Wyatt et al., 2013).

1.2.2.3. SLX-MUS complex

It has been shown that MUS81-EME1 interacts with SLX4. Indeed, these two resolvases participate in the same pathway to generate COs *in vivo* (Castor et al., 2013; Wyatt et al., 2013). The SLX-MUS complex formation is cell cycle dependent. The association is only observed in G_2/M phase, suggesting that the SLX-MUS complex is assembled to eliminate recombinational intermediates before mitosis and cell division. Biochemical studies further confirmed that SLX-MUS has higher efficiency for HJs processing compared to each sub-

complex alone (Wyatt et al., 2013). SLX-MUS uses a "nick and counter-nick mechanism" to resolve HJs with SLX1 making the initial rate-limiting cleavage, and the nicked Holliday junction intermediates is further cleaved by MUS81 on the opposite site across the junction. These two catalytic steps are coordinated and happen within the lifetime of the SLX-MUS-HJ complex. However, the cleavage pattern is asymmetry therefore further processing is required. A more recent study indicated that another heterodimeric nuclease, XPF-ERCC1, could also attach on SLX-MUS and forming the hexameric "SMX tri-nuclease complex", with three catalytic centers (Wyatt et al., 2017). This study extends our knowledge on joint molecules resolution that the scaffold protein SLX4 is the hub for integrating the processing of DNA structures from various repair pathways. Interestingly, XPF-ERCC1 is a known player in nucleotide-excision repair (NER) as well as ICL repair and SLX4 has been recently proposed to be part of the Fanconi anemia complex as FANCP and in this capacity, helping to process ICLs (Cybulski and Howlett, 2011). The mechanisms of how these enzymes coordinate each other and their functions *in vivo* are still elusive and have to be addressed by further studies.

1.3. Classical Holliday junction resolvase GEN1

The search for a classical Holliday junction resolvase in eukaryotes was a challenging task for decades. Even though a RuvC-like activity has been observed from a calf thymus tissue extract and mammalian cells, the corresponding enzyme for this activity was not successfully identified due to low expression levels, lack of sequence homology and the existence of alternative, partially redundant HJ processing pathways (Constantinou et al., 2001; Elborough and West, 1990; Hyde et al., 1994). In 2008, Ip et al eventually demonstrated that human GEN1 and its budding yeast ortholog, Yen1 are the classical HJ resolvases by using two independent approaches (Ip et al., 2008). Human GEN1 was identified by a "brute-force" strategy, in which nuclear extracts from HeLa cells were fractionated through extensive chromatographic purification steps and followed by a series of *in vitro* cleavage tests. In parallel, the activity of about 1100 epitope-tagged proteins from *S. cerevisiae* was assayed individually with model HJ substrates leading to the identification of Yen1 (West, 2009). The followed studies confirmed that the GEN1/Yen1 orthologs function as classical resolvases in almost all model organisms such as *D. melanogaster*, *C. elegans*, *A. thaliana, O. sativa* and *C. thermophilum*, the only exception is *S. pombe,* in which this gene is absent

(Andersen et al., 2011; Bailly et al., 2010; Bauknecht and Kobbe, 2014; Bellendir et al., 2017; Freeman et al., 2014; Wang et al., 2017).

1.3.1. Biochemical properties of GEN1

1.3.1.1. GEN1 belongs to Rad2/XPG nuclease family

Human GEN1 is a polypeptide containing 908 residues. The N-terminal part of the protein (residue 1-389) harbors a Rad2/XPG nuclease core, followed by a long C-terminal region whose sequence is diverse among different species and predicted as disordered **(Figure 1-6)** (Ip et al., 2008; Rass et al., 2010). Members in this family are Mg^{2+}-dependent structure-selective 5′ nucleases that recognize specific DNA secondary structures and show only little or no dependence on nucleotide sequences. The nuclease core is composed by three iconic elements: the XPG N-terminal (XPG-N) and internal (XPG-I) domain and a DNA-binding helix-hairpin-helix (HhH) domain, which is part of a 5′-3′ exonuclease C-terminal domain **(Figure 1-6)**. Based on the similarity, the members can be further classified into four subgroups including FEN1, EXO1, XPG and GEN1, respectively. Each member targets different types of DNA structures and related to important biological functions: FEN1 cuts 5′ flaps or double flaps, which is essential for Okazaki fragment maturation during replication and long-patch base excision repair (BER); EXO1 recognizes 3′ overhangs or nicked duplexes

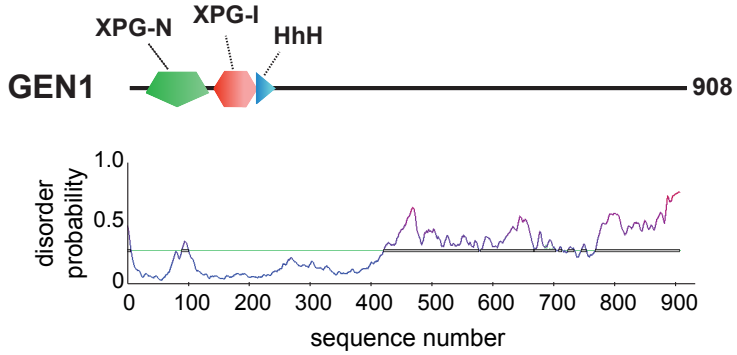


Figure 1-6. Domain architecture of human GEN1.

A Rad2/XPG nuclease core locates at the N-terminal of GEN1, comprising XPG N-terminal (XPG-N), internal (XPG-I) domain and a helix-hairpin-helix (HhH) domain. Sequences at the C-terminal are diverse and unstructured. Disorder prediction is showed below. Figure is adapted from (Rass et al., 2010).

that play an important role in mismatch repair (MMR), telomere maintenance and DNA-end resection during homologous recombination; XPG cleaves bubble structures that are critical for nucleotide-excision repair (NER); and GEN1 is primarily working on Holliday junctions and replication intermediates **(Figure 1-7, panel A)** (Grasby et al., 2012; Ip et al., 2008; Nishino et al., 2006; Tomlinson et al., 2010; Tsutakawa et al., 2014). Despite the diverse substrate specificity, Rad2/XPG nucleases adopted a unified mechanism for the catalysis. These enzymes recognize a bendable DNA structure by binding to ss-ds or ds-ds junctions. The reaction is carried out on the duplex stem by a "two nucleotide-unpairing" mechanism, in which the nuclease unwinds the first two nucleotides from the junction, creating a partially opened intermediate and the subsequent incision is generated at the position one nucleotide away from the junction point **(Figure 1-7, panel B)** (Grasby et al., 2012).

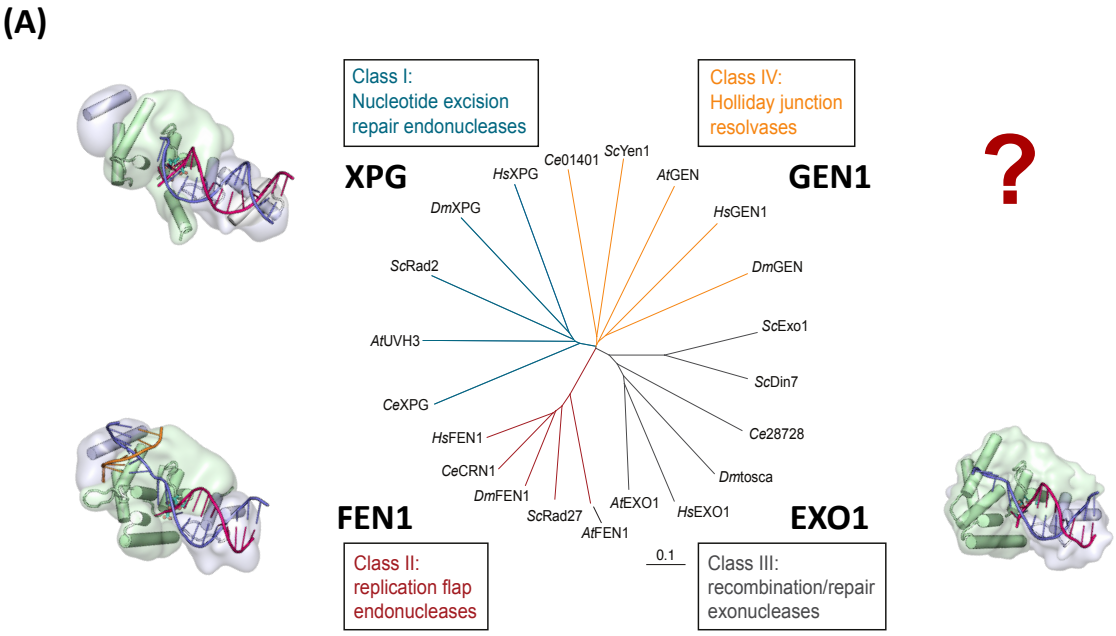


Figure 1-7. Rad2/XPG nuclease family.

(A) Rad2/XPG nucleases can be classified into four subgroups, each of them participate in different DNA repair pathways. XPG is a critical component in nucleotide excision repair (NER) that cleaves the bubble structures. FEN1 is the essential enzyme for Okazaki fragment maturation that removes the double-flap intermediates. EXO1 targeting to 3' overhangs that it is important for mismatch repair and the DNA resection during HR. GEN1 is the resolvase that recognizes HJs. Crystal structures of protein-substrate complexes are presented on the side, except for GEN1, whose structure was not known at the beginning of this work. PDB codes: S. cerevisiae Rad2 (4Q0W), human FEN1 (3Q8K), human EXO1 (3QE9). (B) Rad2/XPG nucleases use a unified mechanism for cleavage but recognize diverse DNA structures. The enzymes bind to a common reacting duplex and make an incision 1 nt away from the junction point. GEN1 is the only member that dimerizes on HJs and promotes symmetric incisions. Nucleolytic incisions are presented as red arrows. The figure is adapted from (Grasby et al., 2012; Ip et al., 2008).

1.3.1.2. Mechanisms of GEN1

GEN1 has a broad spectrum of substrate specificity, cleavage of HJs, 5′ flaps, replication forks, gaps and splayed arms have been reported (Bellendir et al., 2017; Freeman et al., 2014; Ishikawa et al., 2004; Kanai et al., 2007; Rass et al., 2010). Holliday junctions are unique substrates with a symmetrical configuration, therefore requiring two active sites to fully resolve the interlinked structure. Indeed, GEN1 symmetrically cleaves HJs and yields ligatable nicked duplexes. Even though GEN1 is a stable monomeric protein in solution, which is distinct from the well-characterized prokaryotic resolvases that are obligatory homodimers, GEN1 can dimerize on HJs to coordinate the dual incisions (Rass et al., 2010). Moreover, the cleavage is specific to the structure and has very weak sequence preference to double G residues near to a T-rich region (Shah Punatar et al., 2017). The molecular mechanism of GEN1 dimerization remains unclear. This higher ordered arrangement has been observed via electrophoretic mobility shift assays (EMSAs), atomic force microscopy (AFM) and electron microscopy (EM) (Bellendir et al., 2017; Freeman et al., 2014; Rass et al., 2010). However, large-scale reconstitution of the GEN1-HJ complex was not successful due to the intrinsic instability of the protein. Nevertheless, GEN1 dimerization is characterized by the elegant cruciform plasmid assay, which demonstrates that the cleavage of GEN1 on

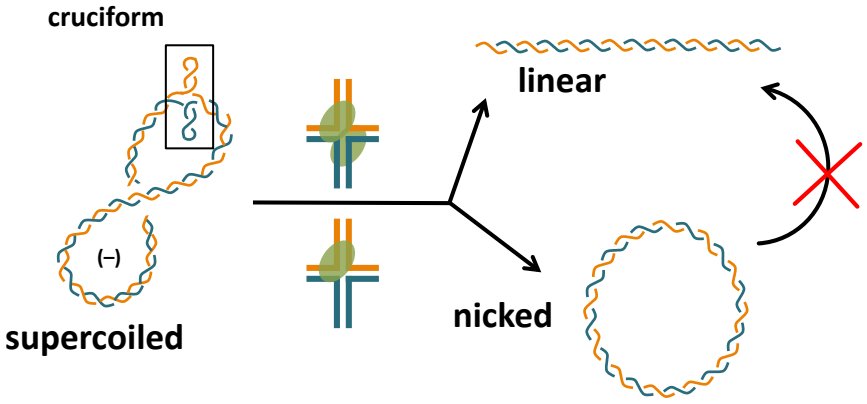


Figure 1-8. Principle of cruciform plasmid assay.

Plasmids harboring an inverted repeat sequence can form cruciform structures, or four-way junctions, when negatively supercoiled, which are the substrates for HJ resolvases. When the enzymes create only one incision, the nicked products are generated and the cruciform structures are reabsorbed, therefore, the second cleavage is not available. If the enzymes promote cooperative dual incisions and the cleavages are within the lifetime of the enzyme-DNA complex, linear products are generated. Supercoiled, nicked and linear DNA can be separated on electrophoresis and as the indication of the enzyme cleavage mode.

HJs uses a "nick and counter-nick" mechanism **(Figure 1-8)**. One of the subunit creates the first, rate-limiting cleavage, and the other subunit introduces the second incision on the diametrically opposite position across the junction. The dual incision is carried out within the lifetime of the GEN1-HJ complex, ensuring that the cleavage is coordinated and prevents unwanted degradation of the substrate (Chan and West, 2015; Lilley and Markham, 1983; Rass et al., 2010).

1.3.2. GEN1 functions in genome maintenance

1.3.2.1. Biological significance of GEN1

The evidence of GEN1 resolving HJs *in vivo* is supported by a genetic study that ectopic expression of human GEN1$^{1\text{-}527}$ in *mus81Δ S. pombe* (in the absence of the functional GEN1 ortholog) rescues meiotic phenotypes (Lorenz et al., 2010). A further delineation of the contribution of each resolvase has been challenging to study due to the fact that many of them play redundant roles and the existence of alternative pathways. Depleting anyone of the resolvases in *S. cerevisiae* often does not show any significant phenotype. A *mus81* knock-out strain shows reduced viability in budding yeast, when treated with DNA damaging agents, but *yen1Δ* cells have no effect. However, the *mus81Δ*, *yen1Δ* double-mutation further exacerbates the sensitivity to DNA damage agents suggesting that Mus81 is the dominant resolvase and Yen1 serves as a backup system (Blanco et al., 2010; Tay and Wu, 2010). However, the preference of resolvases is species-dependent, GEN in *D. melanogaster* and GEN-1 in *C. elegans* are the predominant resolvases for removing recombination intermediates over MUS81 and SLX4 orthologs (Andersen et al., 2011; Bailly et al., 2010).

In human cells, transient depletion of GEN1+SLX4 or GEN1+MUS81 causes anaphase bridges and lagging chromosomes, which are the indications of the existence of unresolved joint molecules. Elevated levels of micronuclei and multi-nucleation are observed in these cells, demonstrating that these resolvases are essential for genome stability (Sarbajna et al., 2014). Other mitotic defects like elongated and segmented chromosomes are also observed in the SLX4-deficient cells with BLM- and GEN1-depletion (Garner et al., 2013). Moreover, knockdown of GEN1 or SLX-MUS reduces the level of SCEs in BLM-null cells derived from Bloom's syndrome patient, indicating that GEN1 and SLX-MUS independently contribute to the COs formation during recombination (Wechsler et al., 2011; Wyatt et al., 2013).

1.3.2.2. Regulation of resolvases in the cell

To avoid sister chromatid exchanges (SCEs) and loss of heterozygosity (LOH), dividing cells primarily use the dissolution pathway to remove HJs and generate NCO products. Indeed, the genes related to dissolution are activated throughout the cell cycle. However, the remaining recombination intermediates that escaped from dissolution, and the intermediates which cannot be processed by the BTR/STR complex (such as single HJs and nicked HJs) are removed by resolvases (Blanco and Matos, 2015; Dehe and Gaillard, 2017; Matos and West, 2014; Wild and Matos, 2016). It has been shown that the resolution pathways are timely and spatially controlled during the cell cycle **(Figure 1-9)**. In *S. cerevisiae*, the activity of Mus81-Mms4 (the ortholog of human MUS81-EME1) is up-regulated in M phase by the kinases Cdc28 and Cdc5, the orthologs of human CDK1 and PLK1, respectively (Gallo-Fernandez et al., 2012; Matos et al., 2011; Szakal and Branzei, 2013). On the other hand, Cdc28 inhibits Yen1. The phosphorylation does not only reduce the enzyme activity but also controls the

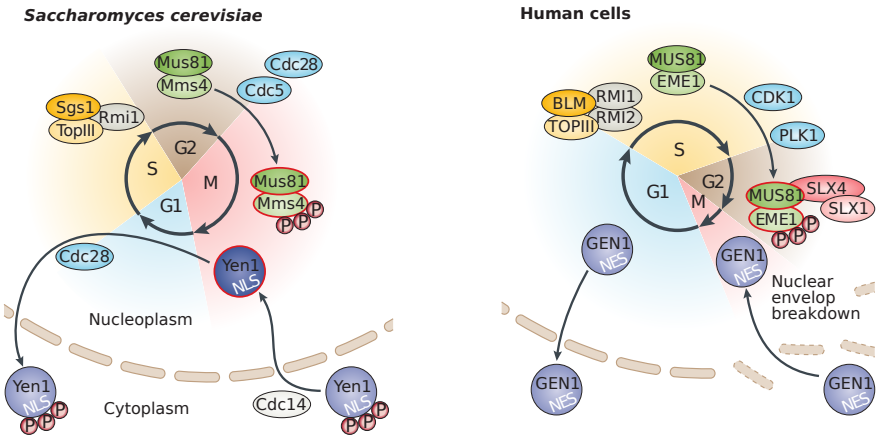


Figure 1-9. Cell cycle regulation of eukaryotic resolvases.

In S. cerevisiae, Mus81-Mms4 is hyperactivated in G_2/M transition by the cell cycle kinases Cdc28 and Cdc5. While phopsphorylated Yen1 is inhibited and remains in the cytoplasm. During anaphase, the phosphatase Cdc14 activates Yen1 by up-regulating the enzyme activity and releasing the nuclear localization signal (NLS). In human cells, EME1 is phosphorylated by CDK1 and PLK1 in G_2/M phase that promoting the SLX1-SLX4-MUS81-EME1 complex formation. On the other hand, GEN1 harbors a nuclear export signal (NES), therefore is excluded from the nucleus. GEN1 gains access to the DNA when the nuclear membrane breakdown in late stage of mitosis. The figure is adapted from (Dehe and Gaillard, 2017).

subcellular localization. Cdc28 phosphorylates the nuclear localization signal (NLS) on Yen1, therefore keeps the enzyme in the cytoplasm. The enzyme activity is later restored and the NLS is released at anaphase by the phosphatase Cdc14 (Blanco et al., 2014; Eissler et al., 2014; Garcia-Luis et al., 2014). In human cells, MUS81-EME1 is controlled by CDK and PLK1 phosphorylation. Specifically, phosphorylated EME1 enhances the SLX-MUS complex formation, which promotes the coordinated processing on HJs (Castor et al., 2013; Wyatt et al., 2013). Distinct from Yen1, human GEN1 is not regulated by phosphorylation events. GEN1 is excluded from the nucleus by its nuclear export signal (NES) in the C-terminal part of the protein. The enzyme only gains access to the genomic DNA when the nuclear membrane breakdown at a late stage of mitosis (Chan and West, 2014).

Taken together, cells adopted a complicated regulation system during evolution along the cell cycle to assure the elimination of joint molecules and minimize the danger of SCEs and LOH. During S phase, dissolution is the predominant pathway to remove joint molecules from replication and recombination, and the activity of SLX-MUS and GEN1 are restrained to prevent the competition of dissolution. At G_2/M transition, a first wave of resolution, SLX-MUS, is activated to resolve the remaining joint molecules. Finally, at anaphase a second wave of resolution, GEN1, serves as the last safeguard to clean up all the joint molecules escaped from previous pathways before chromosome segregation.

1.4. Aim of the thesis

As the classical Holliday junction resolvase in eukaryotes, GEN1 safeguards genome integrity by eliminating all unresolved joint molecules before the chromosome segregation. Genetic and biochemical studies have uncovered the general features of GEN1 regulation and its biological functions. However, the mechanism of GEN1 catalysis remains elusive, especially at the atomic level. Interestingly, GEN1 shares no sequence homology to the well characterized HJ resolvases in prokaryotes and bacteriophages that target to the same substrate. In addition, GEN1 is the sole member in the Rad2/XPG nuclease family that can recognize and dimerize on four-way junctions. It is an open question how GEN1 discriminates different DNA structures by using the conserved nuclease core. Therefore, studying GEN1 does not only provide knowledge about the convergent evolution of HJ recognition, but also it highlights the divergent evolution of substrate recognition by Rad2/XPG nucleases.

The aim of this thesis focused on the structural and biochemical characterization of human GEN1. The goal is to reveal the molecular basis of GEN1-DNA interactions by structural biology techniques in combination with biophysical and biochemical tools to understand substrate recognition, protein dimerization and functional features of GEN1.

Chapter 2 –

Structure of Human GEN1-DNA Complex: The Chromodomain is required for Efficient Substrate Recognition and Cleavage

Shun-Hsiao Lee, Lissa Nicola Princz, Maren Felizitas Klügel, Bianca Habermann, Boris Pfander, Christian Biertümpfel

Human Holliday junction resolvase GEN1 uses a chromodomain for efficient DNA recognition and cleavage. eLife 4:e12256 (2015)

This study presents the first crystal structure of human GEN1 HJ resolvase. The architecture reveals an iconic XPG nuclease core appended by a chromodomain, which interacts with DNA. The chromodomain has critical functions in substrate recognition and cleavage as demonstrated by *in vitro* and *in vivo* experiments.

This study was conducted under the supervision of Dr. Christian Biertümpfel. Lee, S.-H. participated in the conception and design of the project, data acquisition, data analysis and interpretation and writing the manuscript. The *in vivo* experiments and part of the bioinformatic analysis were contributed by collaborators. Detailed author contributions are included in the attached article.

eLIFE elifesciences.org

RESEARCH ARTICLE

Human Holliday junction resolvase GEN1 uses a chromodomain for efficient DNA recognition and cleavage

Shun-Hsiao Lee[1], Lissa Nicola Princz[2], Maren Felizitas Klügel[1], Bianca Habermann[3], Boris Pfander[2], Christian Biertümpfel[1]*

[1]Department of Structural Cell Biology, Molecular Mechanisms of DNA Repair, Max Planck Institute of Biochemistry, Martinsried, Germany; [2]Department of Molecular Cell Biology, DNA Replication and Genome Integrity, Max Planck Institute of Biochemistry, Martinsried, Germany; [3]Computational Biology, Max Planck Institute of Biochemistry, Martinsried, Germany

Abstract Holliday junctions (HJs) are key DNA intermediates in homologous recombination. They link homologous DNA strands and have to be faithfully removed for proper DNA segregation and genome integrity. Here, we present the crystal structure of human HJ resolvase GEN1 complexed with DNA at 3.0 Å resolution. The GEN1 core is similar to other Rad2/XPG nucleases. However, unlike other members of the superfamily, GEN1 contains a chromodomain as an additional DNA interaction site. Chromodomains are known for their chromatin-targeting function in chromatin remodelers and histone(de)acetylases but they have not previously been found in nucleases. The GEN1 chromodomain directly contacts DNA and its truncation severely hampers GEN1's catalytic activity. Structure-guided mutations in vitro and in vivo in yeast validated our mechanistic findings. Our study provides the missing structure in the Rad2/XPG family and insights how a well-conserved nuclease core acquires versatility in recognizing diverse substrates for DNA repair and maintenance.

DOI: 10.7554/eLife.12256.001

***For correspondence:** biertuempfel@biochem.mpg.de

Competing interests: The authors declare that no competing interests exist.

Funding: See page 20

Received: 12 October 2015
Accepted: 17 December 2015
Published: 18 December 2015

Reviewing editor: Volker Dötsch, Goethe University, Germany

Introduction

Homologous recombination (HR) is a fundamental pathway ensuring genome integrity and genetic variability (***Heyer, 2015***). In mitotic cells, double-strand breaks (DSBs) can be repaired by HR using the sister chromatid as a template to restore the information in the complementary double strand. In meiosis, the repair of programmed DSBs by HR and the formation of crossovers are crucial to provide physical linkages between homologs and to segregate homologous chromosomes. Furthermore, HR during meiosis creates sequence diversity in the offspring through the exchange between homologs (***Petronczki et al., 2003***; ***Sarbajna and West, 2014***).

HR proceeds by pathways that may lead to the formation of DNA four-way junctions or Holliday junctions (HJs) that physically link two homologous DNA duplexes (***Heyer, 2015***; ***Holliday, 1964***; ***Schwacha and Kleckner, 1995***; ***Szostak et al., 1983***). Faithful removal of HJs is critical to avoid chromosome aberrations (***Wechsler et al., 2011***) and cells have evolved sophisticated measures to disentangle joint molecules. One basic mechanism is resolution mediated by HJ resolvases that introduce precise symmetrical nicks into the DNA at the branch point. Nicked DNA strands are then rejoined by endogenous ligases leading to fully restored or recombined DNA strands. This mechanism is well studied for bacterial and bacteriophage resolvases such as *Escherichia coli* RuvC, T7 endonuclease I, T4 endonuclease VII (***Benson and West, 1994***; ***Lilley and White, 2001***). These resolvases operate as dimers and show a large degree of conformational flexibility in substrate

eLife digest Factors like ultraviolet radiation and harmful chemicals can damage DNA inside living cells, which can lead to breaks that form across both strands in the DNA double helix. "Homologous recombination" is one of the major mechanisms by which cells repair these double-strand breaks. During this process, the broken DNA interacts with another undamaged copy of the DNA to form a special four-way structure called a "Holliday junction". The intact DNA strands are then used as templates to repair the broken strands. However, once this has occurred the Holliday junction needs to be 'resolved' so that the DNA strands can disentangle.

One way in which Holliday junctions are resolved is through the introduction of precise symmetrical cuts in the DNA at the junction by an enzyme that acts like a pair of molecular scissors. Re-joining these cut strands then fully restores the DNA. Enzymes that generate the cuts in DNA are called nucleases, and the nuclease GEN1 is crucial for resolving Holliday junctions in organisms such as fungi, plants and animals. GEN1 belongs to a family of enzymes that act on various types of DNA structures that are formed either during damage repair, DNA duplication or cell division. However, GEN1 is the only enzyme in the family that can also recognize a Holliday junction and it was unclear why this might be.

Lee et al. have now used a technique called X-ray crystallography to solve the three-dimensional structure of the human version of GEN1 bound to a Holliday junction. This analysis revealed that many features in GEN1 resemble those found in other members of the same nuclease family. These features include two surfaces of the protein that bind to DNA and are separated by a wedge, which introduces a sharp bend in the DNA. However, Lee et al. also found that GEN1 contains an additional region known as a "chromodomain" that further anchors the enzyme to the DNA. The chromodomain allows GEN1 to correctly position itself against DNA molecules, and without the chromodomain, GEN1's ability to cut DNA in a test tube was severely impaired. Further experiments showed that the chromodomain was also important for GEN1's activity in yeast cells growing under stressed conditions.

The discovery of a chromodomain in this human nuclease may provide many new insights into how GEN1 is regulated, and further work could investigate if this chromodomain is also involved in binding to other proteins.

DOI: 10.7554/eLife.12256.002

recognition and in aligning both active sites for coordinated cleavage. Interestingly, T4 endonuclease VII and RuvC reach into and widen the DNA junction point whereas T7 endonuclease I binds DNA by embracing HJs at the branch point (***Biertümpfel et al., 2007***; ***Górecka et al., 2013***; ***Hadden et al., 2007***).

In eukaryotes, HR is more complex and tightly regulated. In somatic cells, HJ dissolution by a combined action of a helicase and a topoisomerase (BLM-TOPIIIα-RMI1-RMI2 complex in humans) is generally the favored pathway, possibly to restore the original (non-crossover) DNA arrangement (***Cejka et al., 2010***, ***2012***; ***Ira et al., 2003***; ***Putnam et al., 2009***; ***Wu and Hickson, 2003***). In contrast, HJ resolution generates crossover and non-crossover arrangements depending on cleavage direction. Several endonucleases such as GEN1, MUS81-EME1, and SLX1-SLX4 have been implicated as HJ resolvases in eukaryotes (***Andersen et al., 2011***; ***Castor et al., 2013***; ***Fekairi et al., 2009***; ***Garner et al., 2013***; ***Ip et al., 2008***; ***Muñoz et al., 2009***; ***Svendsen and Harper, 2010***; ***Svendsen et al., 2009***; ***Wyatt et al., 2013***). Interestingly, these resolvases are not structurally related and have different domain architectures, giving rise to variable DNA recognition and regulation mechanisms. The interplay between resolution and dissolution mechanisms is not fully understood yet, however, cell cycle regulation of resolvases seems to play an important role (***Blanco et al., 2014***; ***Chan and West, 2014***; ***Eissler et al., 2014***; ***Matos et al., 2011***).

GEN1 belongs to the Rad2/XPG family of structure-selective nucleases that are conserved from yeast to humans (***Ip et al., 2008***; ***Lieber, 1997***; ***Yang, 2011***). The Rad2/XPG family has four members with different substrate preferences that function in DNA maintenance (***Nishino et al., 2006***; ***Tsutakawa et al., 2014***). They share a conserved N-terminal domain (XPG-N), an internal domain (XPG-I) and a 5'->3' exonuclease C-terminal domain containing a conserved helix-hairpin-helix motif.

C-terminal to the nuclease core is a regulatory region that is diverse in sequence and predicted to be largely unstructured. Although the catalytic cores are well conserved in the superfamily, substrate recognition is highly diverse: XPG/Rad2/ERCC5 recognizes bubble/loop structures during nucleotide-excision repair (NER), FEN1 cleaves flap substrates during Okazaki fragment processing in DNA replication, EXO1 is a 5'->3' exonuclease that is involved in HR and DNA mismatch repair (MMR) and GEN1 recognizes Holliday junctions (***Grasby et al., 2012***; ***Ip et al., 2008***; ***Nishino et al., 2006***; ***Tomlinson et al., 2010***; ***Tsutakawa et al., 2014***). A common feature of the superfamily is their inherent ability to recognize flexible or bendable regions in the normally rather stiff DNA double helix. Interestingly, GEN1 shows versatile substrate recognition accommodating 5' flaps, gaps, replication fork intermediates and Holliday junctions (***Ip et al., 2008***; ***Ishikawa et al., 2004***; ***Kanai et al., 2007***). According to the current model, however, the primary function of GEN1 is HJ resolution (***Garner et al., 2013***; ***Sarbajna and West, 2014***; ***West et al., 2015***) and it is suggested to be a last resort for the removal of joint molecules before cytokinesis (***Matos et al., 2011***).

To date, structural information is available for all members of the family but GEN1 (***Mietus et al., 2014***; ***Orans et al., 2011***; ***Tsutakawa et al., 2011***). A unified feature of these structures is the presence of two DNA-binding interfaces separated by a hydrophobic wedge. This wedge is composed of two protruding helices that induce a sharp bend into flexible DNA substrates. Rad2/XPG family members also share a helix-two-turn-helix (H2TH) motif that binds and stabilizes the uncleaved DNA strand downstream of the catalytic center. However, the comparison of DNA recognition features within the Rad2/XPG family has been hampered because of the lack of structural information on GEN1.

To understand the molecular basis of GEN1's substrate recognition, we determined the crystal structure of human GEN1 in complex with HJ DNA. In combination with mutational and functional analysis using in vitro DNA cleavage assays and in vivo survival assays with mutant yeast strains, we highlight GEN1's sophisticated DNA recognition mechanism. We found that GEN1 does not only have the classical DNA recognition features of Rad2/XPG nucleases, but also contains an additional DNA interaction site mediated by a chromodomain. In the absence of the chromodomain, GEN1's catalytic activity was severely impaired. This is the first example showing the direct involvement of a chromodomain in a nuclease. Our structural analysis gives implications for a safety mechanism using an adjustable hatch for substrate discrimination and to ensure coordinated and precise cleavage of Holliday junctions.

Results

Structure determination and architecture of the GEN1-DNA complex

In order to structurally characterize human GEN1, we crystallized the catalytically inactive variant $GEN1^{2\text{-}505\ D30N}$, denoted GEN1 for simplicity, in complex with an immobile Holliday junction having arm lengths of 10 bp (***Figure 1***). The structure was determined experimentally and refined up to 3.0 Å resolution with an R_{free} of 0.25 (***Table 1***). The HJ crystallized bridging between two protein monomers in the asymmetric unit (***Figure 1—figure supplement 1***). The overall structure of GEN1 resembles the shape of a downwards-pointing right hand with a 'thumb' extending out from the 'palm' and the DNA is packed against the ball of the thumb (***Figure 1***). The palm contains the catalytic core, which is formed by intertwined XPG-N and XPG-I domains (***Figure 1A/B***, green). They consist of a seven-stranded β-sheet in the center surrounded by nine helices harboring the conserved active site (***Figure 1B/D***, orange). The catalytic residues form a cluster of negatively charged residues (D30, E75, E134, E136, D155, D157, D208) that were originally identified by mutational analysis (***Ip et al., 2008***; ***Lee et al., 2002***; ***Wakasugi et al., 1997***) and are conserved in other Rad2/XPG family members (***Figure 1B/C*** and ***Figure 2***). The XPG-I domain is followed by a 5'->3' exonuclease C-terminal domain (EXO; ***Figure 1B/D***, blue). The EXO domain consists of a helix-two-turn-helix (H2TH) motif (helices α10-α11) accompanied by several α-hairpins (α12-α13 and α14-α15). A similar arrangement is also found in other proteins, which use a H2TH motif for non-sequence specific DNA recognition (***Tomlinson et al., 2010***). The EXO domain in GEN1 has a 78 amino acid insertion (residues 245–322), of which only helix α12b (residues 308–322) is ordered in the structure (***Figure 1A***, gray and ***Figure 2***). Helix α12b packs loosely with the H2TH helices (α10-α11) and helix α12 at the 'finger' part of GEN1. Yeast Rad2, a homolog of human XPG, also contains helix α12b,

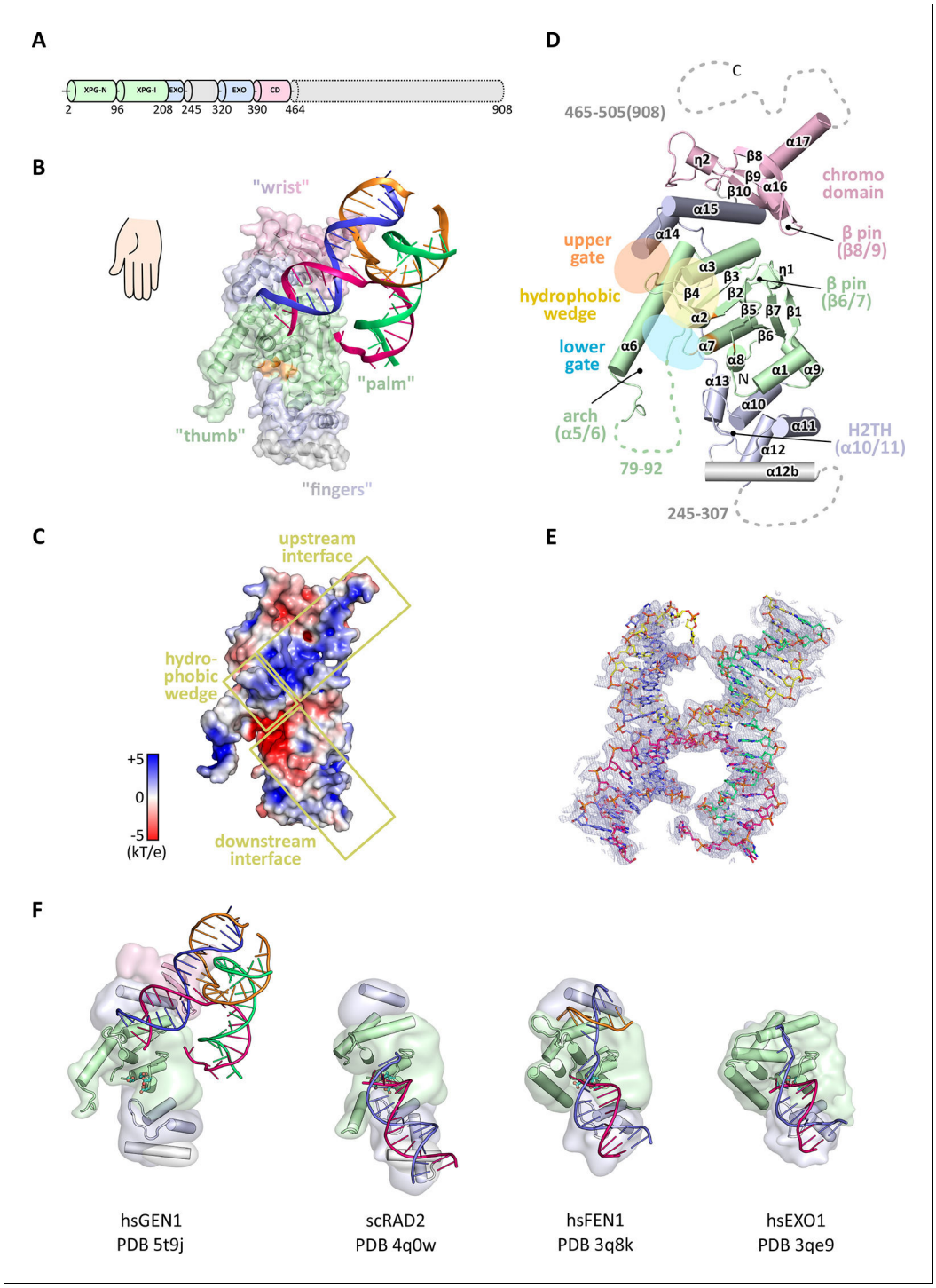


Figure 1. Architecture of human GEN1. (**A**) Domain architecture of human GEN1. The structurally unknown regulatory domain (residues 465–908) is shown with dotted lines. (**B**) Overview of the catalytic core of GEN1 in complex with HJ DNA. The protein resembles the shape of a downwards-pointing right hand with helix α6 as the thumb. The protein is depicted in half transparent surface representation with secondary structure elements underneath. The DNA is shown in ladder representation with individual strands in different colors. The coloring of GEN1 follows domain boundaries: intertwining XPG-N and XPG-I in green, 5′->3′ exonuclease C-terminal domain (EXO) in blue, chromodomain in pink, unassigned regions in gray. Active site residues (E134, E136, D155, D157) are highlighted in orange. (**C**) Electrostatic surface potential of GEN1. The coloring follows the potential from -5 (red) to +5 kT/e (blue). The DNA-binding interfaces and the position of the hydrophobic wedge are marked in yellow. (**D**) Secondary structure elements of the catalytic core of GEN1 in cartoon representation with the same

Figure 1 continued on next page

Figure 1 continued

colors as before. Dotted lines represent parts that are not resolved in the crystal structure. The numbering follows a unified scheme for the Rad2/XPG family (compare ***Figure 2***) for α-helices, β-sheets and 3_{10}-helices (η). (**E**) Experimental electron density map (autoSHARP, solvent flattened, contoured at 1σ) drawn around the HJ in the GEN1 complex. The DNA model is shown in ball-stick representation with carbon atoms of individual strands in different colors (yellow, light blue, magenta, green) and oxygen atoms in red, phosphor atoms in orange, nitrogen atoms in dark blue. (**F**) Structural comparison of Rad2/XPG family nucleases. Proteins are shown in a simplified surface representation with important structural elements in cartoon representation and DNA in ladder representation. The color scheme is the same as in **B**. ***Figure 1—figure supplement 1*** shows the content of the asymmetric unit.

DOI: 10.7554/eLife.12256.003

The following figure supplement is available for figure 1:

Figure supplement 1. Content of the asymmetric unit of the GEN1-HJ crystal.

DOI: 10.7554/eLife.12256.004

and it shows a similar arrangement as in GEN1 (***Figure 1F***). The EXO domain sandwiches the XPG-N/I domains with a long linker reaching from the bottom 'fingers' (α10-α13) along the backside of GEN1 to the top of the XPG-N/I domains at the 'wrist' (α14-α15). A structure-based sequence alignment of the nuclease core of human GEN1, FEN1, EXO1 and yeast Rad2 proteins with functional annotations relates sequence conservation to features in the Rad2/XPG family (***Figure 2***). The comparison with members in the Rad2/XPG identified two DNA binding interfaces and a hydrophobic wedge (ball of the thumb) that separates the upstream and the downstream interface (***Figure 1C/D*** and compare ***Figure 1F***). GEN1 has two prominent grooves close to the hydrophobic wedge, which we termed upper and lower gate or gateway for comparison (***Figure 1D***, orange and blue ellipses, respectively).

Notably, a small globular domain (residues 390–464) was found extending the GEN1 nuclease core at the wrist (***Figure 1***, pink). A DALI search (***Holm and Rosenström, 2010***) against the Protein Data Bank (PDB) identified this domain as a chromodomain (chromatin organization modifier domain). The domain has a chalice-shaped structure with three antiparallel β-strands packed against a C-terminal α-helix and it forms a characteristic aromatic cage. The opening of the chalice abuts helix α15 from the EXO domain.

GEN1 has a conserved chromodomain with a closed aromatic cage

Chromodomains are found in many chromatin-associated proteins that bind modified histone tails for chromatin targeting (reviewed in ***Blus et al., 2011***; ***Eissenberg, 2012***; ***Yap and Zhou, 2011***), but it has not previously been associated with nucleases. To understand the significance of the chromodomain for the function of GEN1, we first examined if the chromodomain is conserved in GEN1 homologs using HMM-HMM (Hidden Markov Models) comparisons in HHPRED (***Söding et al., 2005***). We found that the chromodomain in GEN1 is conserved from yeast (Yen1) to humans (***Figure 3A***). The only exception is *Caenorhabditis elegans* GEN1, which has a much smaller protein size of 443 amino acids compared to yeast Yen1 (759 aa) or human GEN1 (908 aa).

To further compare the structural arrangement of the aromatic cage in human GEN1 with other chromodomains, we analyzed the best matches from the DALI search (***Figure 3B***). We found many hits for different chromo- and chromo-shadow domains with root mean square deviations between 1.9 and 2.8 Å (compare ***Figure 3—source data 1***). A superposition of the aromatic cage of the five structurally most similar proteins with GEN1 (***Figure 3C***) showed that residues W418, T438, and E440 are well conserved, whereas two residues at the rim of the canonical binding cleft are changed from phenylalanine/tyrosine to a leucine (L397) in one case and a proline (P421) in another (***Figure 3C***). Instead, Y424 occupies the space proximal to P421, which is about 1.5 Å outwards of the canonical cage and widens the GEN1 cage slightly. The substitution of phenylalanine/tyrosine to leucine is also found in CBX chromo-shadow domains (see below); however, the rest of the GEN1 aromatic cage resembles rather chromodomains.

Chromodomains often recognize modified lysines through their aromatic cage thus targeting proteins to chromatin (reviewed in ***Blus et al., 2011***; ***Eissenberg, 2012***; ***Yap and Zhou, 2011***). Given the conserved aromatic cage in GEN1, we tested the binding to modified histone tail peptides

Table 1. Data collection and refinement statistics.

Data Set	G505-4w006native	G505-4w006Ta peak	G505-4w006SeMet peak
Diffraction Data Statistics			
Synchrotron Beamline	SLS PXII	SLS PXII	SLS PXII
Wavelength	0.99995	1.25473	0.97894
Resolution (Å)	75-3.0	75.4-3.8	43.6-4.4
Space Group	P 3_2	P 3_2	P 3_2
Cell dimensions			
a (Å)	86.94	87.06	87.11
b (Å)	86.94	87.06	87.11
c (Å)	200.72	201.30	199.69
α (°)	90	90	90
β (°)	90	90	90
γ (°)	120	120	120
$I/\sigma I$*	18.4 (1.9)	27.49 (5.83)	16.58 (3.82)
Completeness (%)*	99.8 (98.8)	99.6 (97.3)	97.3 (83.3)
Redundancy*	6.3	10.2	5.1
Rsym (%)*	6.2 (90.7)	7.7 (42.2)	6.9 (43.4)
Refinement Statistics			
Resolution (Å)	75-3.0		
Number of Reflections	33933		
R_{work}/R_{free}	0.199/0.241		
Number of Atoms			
Protein	6298		
DNA	1589		
Water/Solutes	27		
B-factors			
Protein	123.4		
DNA	150.2		
Water/Solutes	92.6		
R.M.S Deviations			
Bond lengths (Å)	0.010		
Bond Angles (°)	0.623		
Ramachandran Plot			
Preferred	753 (97.9 %)		
Allowed	16 (2.1%)		

*Values for the highest resolution shell are shown in parenthesis

DOI: 10.7554/eLife.12256.005

(***Figure 3C/D***). However, we did not detect any binding despite extensive efforts using various histone tail peptides in pull-down assays, microscale-thermophoresis (MST) or fluorescence anisotropy measurements (compare ***Figure 3—source data 2*** and ***Figure 3—figure supplement 2***). Our structure shows that the aromatic cage is closed by helix α15 (***Figure 3E*** blue/pink), which has a hydrophobic interface towards the aromatic cage with residues L376, T380, and M384 reaching into it (compare ***Figure 4F***). This potentially hampers the binding of the tested peptides in this conformation under physiological conditions.

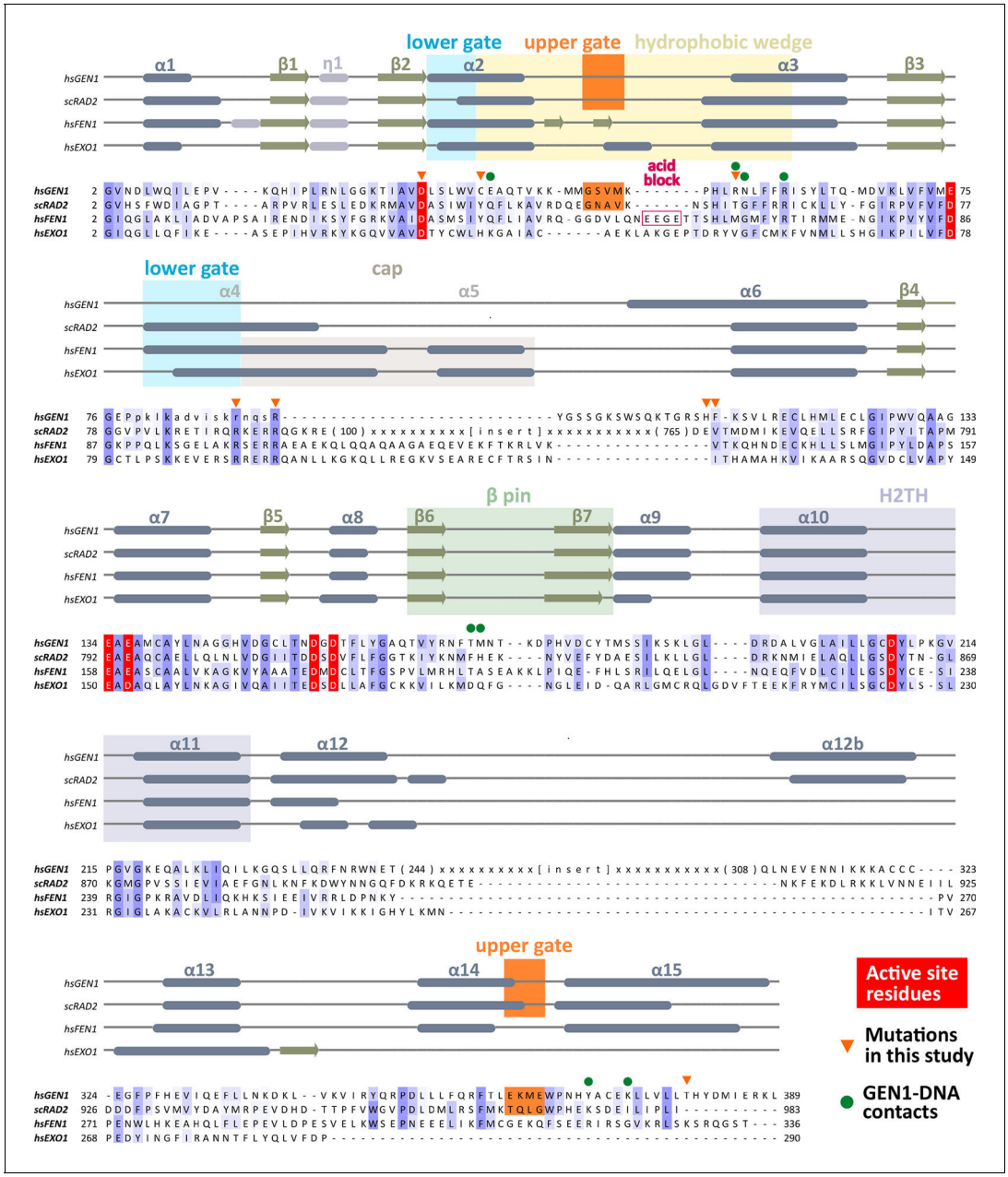


Figure 2. Alignment of the nuclease cores of Rad2/XPG-family proteins. The alignment is based on known crystal structures: human GEN1 (PDB 5t9j, this study), yeast Rad2 (PDB 4q0w), human FEN1 (PDB 3q8k), human EXO1 (3qe9). Secondary structure elements are depicted on top of the sequence with dark blue bars for α-helices, light blue bars for 3_{10}-helices and green arrows for β-sheets. The numbering follows a unified scheme for the superfamily. Functional elements are labeled and described in the main text. Sequences are colored by similarity (BLOSUM62 score) and active site residues are marked in red. Mutations analyzed in this study are marked with an orange triangle and DNA contacts found in the human GEN1–HJ structure have a dark green dot. Disordered or missing parts in the structures are labeled in small letters or with x.

DOI: 10.7554/eLife.12256.006

The GEN1 chromodomain is distantly related to CBX and CDY chromodomains

To explore the functional role of the GEN1 chromodomain, we evaluated its similarity to other chromodomains by comparing all of the 46 known human chromodomains from 34 different proteins. We made pairwise comparisons with HHPRED, PSIBLAST, combined the alignments and generated a phylogenetic tree (***Figure 3F*** and ***Figure 3—figure supplement 1***). The analysis showed a tree

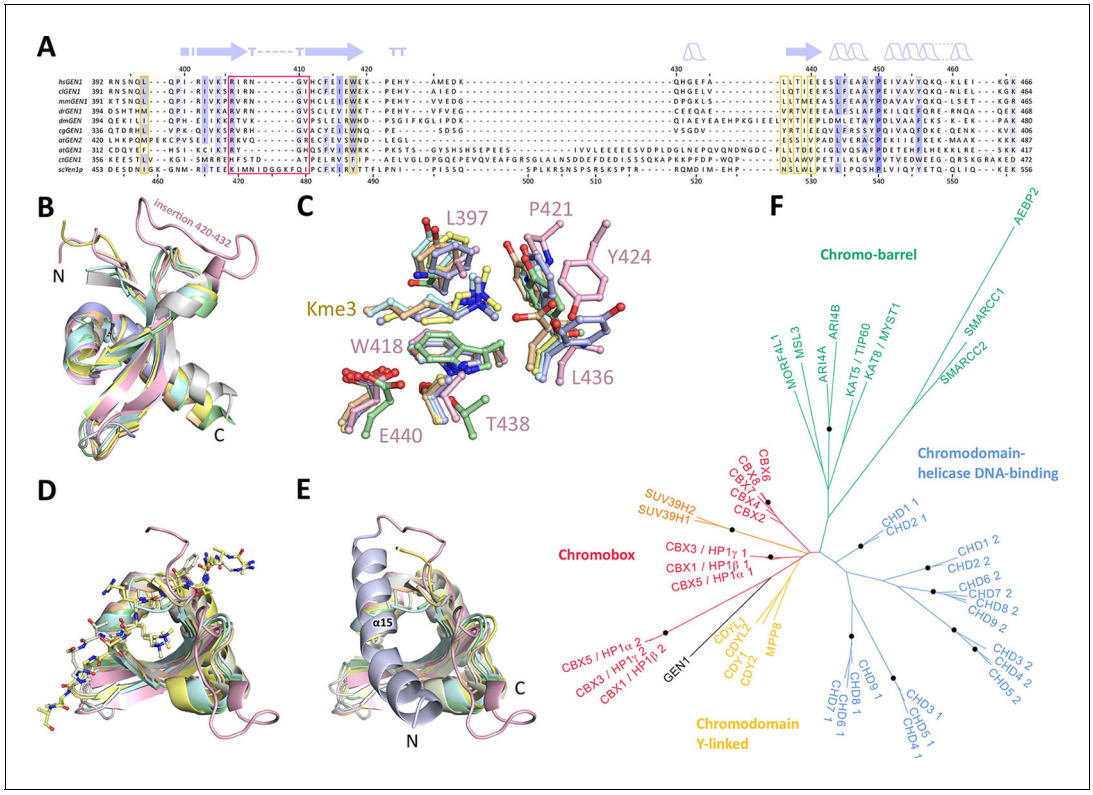


Figure 3. Chromodomain comparison. (**A**) Sequence alignment of GEN1 chromodomains from different organisms: hsGEN1 (*Homo sapiens*), clGEN1 (*Canis lupus*), mmGEN1 (*Mus musculus*), drGEN1 (*Danio rerio*), atGEN1/2 (*Arabidopsis thaliana*), cgGEN1 (*Crassostrea gigas*), scYEN1 (*Saccharomyces cerevisiae*). The presence of a chromodomain is conserved from yeast to human with *Caenorhabditis elegans* as an exception. Secondary structure elements of the GEN1 chromodomain are shown on top. The sequence coloring is based on a similarity matrix (BLOSUM62). The corresponding positions of the DNA-interaction site in human GEN1 is marked with a red box and residues of the aromatic cage are highlighted with a yellow box. (**B**) GEN1 has a canonical chromodomain fold of three antiparallel beta-sheets packed against an α-helix. (**C**) The arrangement of the aromatic cage in GEN1 is comparable to other chromodomains but less aromatic and slightly larger. (**D**) The superposition of different chromodomains places cognate binding peptides of hsMPP8 and mmCBX7 (and others) into the aromatic cage. (**E**) The aromatic cage of GEN1 is closed by helix α15. Panels **B–D** show the chromodomains of hsGEN1 (pink, PDB 5t9j), hsCBX3 (gray, PDB 3kup) hsSUV39H1 (green, PDB 3mts), hsMPP8 (yellow, PDB 3lwe), dmHP1a (orange, chromo shadow PDB 3p7j), dmRHINO (cyan, PDB 4quc/3r93), mmCBX7 (light blue, PDB 4x3s; compare ***Figure 3—source data 1***). (**F**) Phylogenetic tree of all known human chromodomains. GEN1 is distantly related to the CBX chromo-shadow domains and CDY chromodomains. The corresponding alignment for calculating the phylogenetic tree is shown in ***Figure 3—figure supplement 1***. GEN1 is colored in black, chromobox (CBX) proteins are colored in red, interspersed by SUV39H histone acetylases (orange) and chromodomain Y-linked (CDY) proteins (yellow). Chromo-barrel domain proteins are colored in green and chromodomain-helicase DNA-binding (CHD) proteins are in blue. Chromodomains and chromo-shadow domains from the same protein are labeled with 1 and 2, respectively. Stable branches with boostrap values equal or higher than 0.8 are marked with a black dot. The binding of the GEN1 chromodomain to a set of histone peptides was tested but no interaction was detected (***Figure 3—source data 2*** and ***Figure 3—figure supplement 2***).

DOI: 10.7554/eLife.12256.007

The following source data and figure supplements are available for figure 3:

Source data 1. Proteins found in a DALI search.

DOI: 10.7554/eLife.12256.008

Source data 2. N-terminally fluorescein-labeled peptides used for chromodomain binding assays.

DOI: 10.7554/eLife.12256.009

Figure supplement 1. Sequence alignment of all known human chromodomains.

DOI: 10.7554/eLife.12256.010

Figure supplement 2. Histone peptide pull-down assay.

DOI: 10.7554/eLife.12256.011

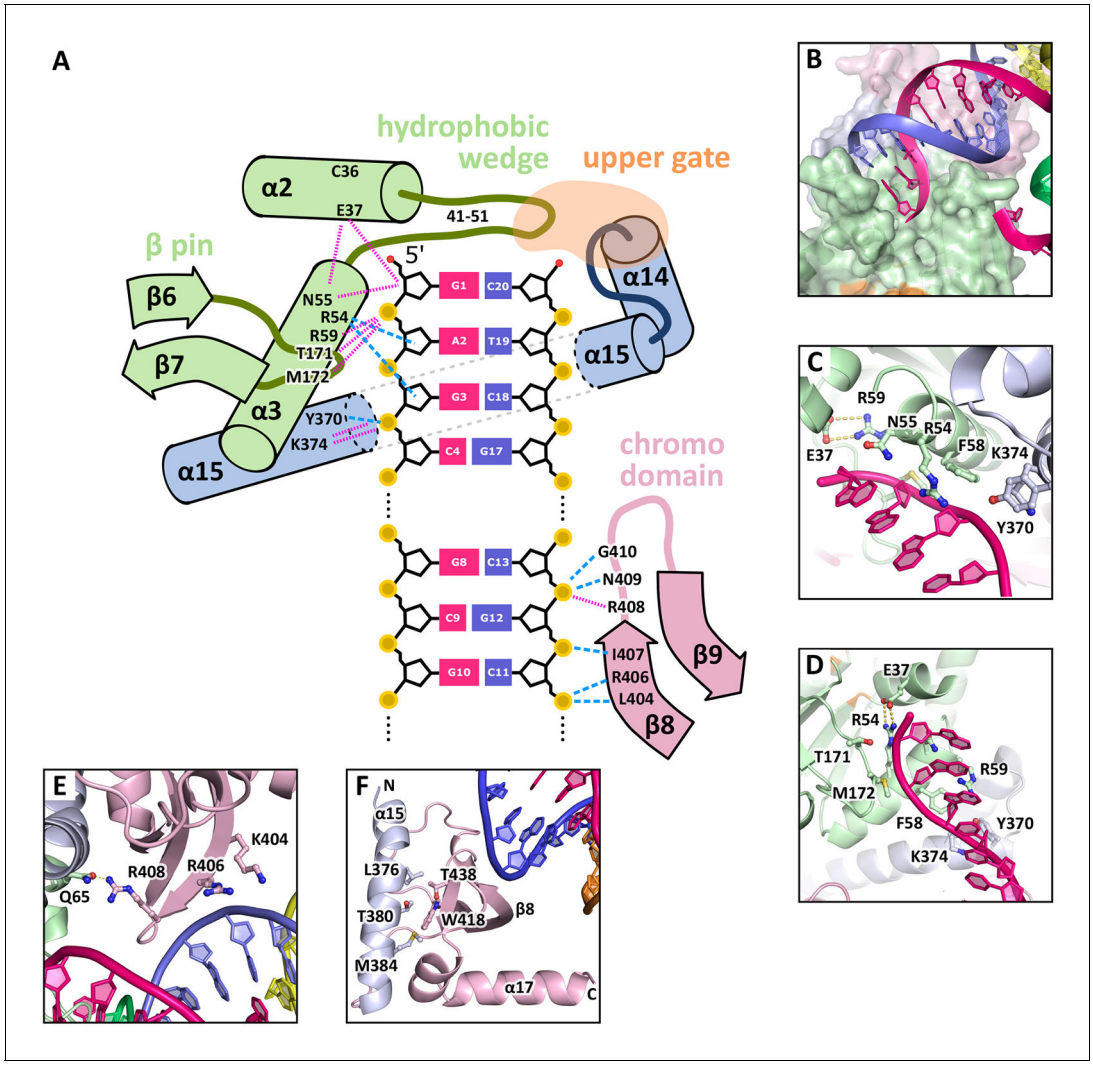


Figure 4. DNA interactions in the GEN1-DNA complex. (**A**) Schematic of the GEN1-DNA interactions at the upstream interface. The coloring is the same as in ***Figure 1***. The nuclease core (green and blue) interacts with the uncleaved strand and the chromodomain (pink) contacts the complementary strand. Hydrogen bonds are shown with blue dashed lines and van-der-Waals contacts are in red dotted lines. (**B**) Interactions at the hydrophobic wedge. The end of the DNA double helix docks onto the hydrophobic wedge formed by helices α2 and α3. (**C/D**) Interactions with the uncleaved strand in two views. All key residues form sequence-independent contacts to the DNA backbone. R54 reaches into the minor groove of the DNA. The complementary DNA strand has been removed for clarity (**E/F**) Interactions of the chromodomain with the complementary strand in two views. The backbone of residues 406–410 (β-hairpin β8-β9) abuts the DNA backbone. R406 has a supporting role in the interaction and R408 forms a polar interaction with Q65, which establishes a connection between the chromodomain and the nuclease core. Helix α15 makes hydrophobic interactions with the aromatic cage and thus blocks it.

DOI: 10.7554/eLife.12256.012

branching into known subfamilies: chromobox proteins (CBX, red), chromodomain Y-linked proteins (CDY, yellow), chromodomain-helicase DNA-binding proteins (blue) and chromo-barrel domain proteins (green). The GEN1 chromodomain was found to be distantly related to the CDY chromodomains and chromobox proteins, particularly to the chromo-shadow domains of CBX1, CBX3 and CBX5. This agrees with the result from the DALI search, in which CBX chromo-shadow domains and homologs thereof were among the closest structural matches. Together with the observed differences in residues forming the aromatic cage, it indicates that the GEN1 chromodomain forms a new subgroup with features from chromo- and chromo-shadow domains that emerged from a common ancestor within CBX/CDY proteins.

GEN1-DNA interactions

The GEN1-HJ structure revealed that the upstream DNA-binding interface acts as a docking site for double-stranded DNA and that the chromodomain secures its position. The DNA is bound at the upstream interface and the hydrophobic wedge but does not extend into the active site or to the downstream interface (***Figure 1B/C/D***). Comparison of the structure of GEN1 to related structures of FEN1, Rad2 and EXO1 (***Miętus et al., 2014***; ***Orans et al., 2011***; ***Tsutakawa et al., 2011***) suggests that a DNA substrate has to extend to the downstream interface to position a DNA strand for cleavage by the active site of GEN1 (***Figure 1B/C*** and ***Figure 1F***). In the GEN1 structure, the end of the DNA arm attaches to the hydrophobic wedge provided by parts of helices α2-α3 and their connecting loop (***Figure 4A/B***), forming van-der-Waals contacts with the first base pair, which docks perfectly onto the protruding curb of residues 41–51 (***Figure 4B***). The uncleaved DNA strand is further stabilized and its geometrical arrangement is fixed by the upstream DNA-binding interface. Particularly, the DNA is contacted by a β-pin (strands β6-β7; ***Figure 4A/C***) from one side and by R54 and F58 (***Figure 4A/D***) from helix α3 together with Y370 and K374 (helix α15) from the opposite side (***Figure 4A/C***). The key residues in the β-pin are T171 that forms a hydrogen bridge to the phosphate of the first base (***Figure 4A***, 'G1') and M172 that makes a van-der-Waals contact to the DNA backbone at the second base (***Figure 4A***, 'A2'). R54 reaches into the DNA minor groove and forms a hydrogen bond with the ribose ring oxygen at the third base of the uncleaved strand and F58 packs against the same ribose moiety (***Figure 4C/D***). Y370 and K374 in α15 form hydrogen bonds to the backbone of the third base of the uncleaved DNA strand (***Figure 4D***, 'G3').

An additional interaction point is provided by a β-hairpin from the chromodomain (strands β8-β9), one DNA turn upstream of the hydrophobic wedge (***Figure 4A/E/F***). This β-hairpin interacts with the complementary DNA strand by matching the protein backbone (residues 406–411) to the contour of the DNA backbone in a sequence unspecific manner (***Figure 4A/E***). The side chains of K404 and R406 project out, and they are in hydrogen bonding distance to the DNA (***Figure 4E***). Remarkably, R408 forms a polar interaction with Q65, which establishes a connection between the DNA contact point at the chromodomain and the nuclease core (***Figure 4E***). The interactions at the chromodomain extend the upstream DNA-binding interface to cover a full DNA turn, reinforcing the binding.

The downstream binding interface can be inferred from other Rad2/XPG structures (***Figure 1C/F***) as the nuclease core is well conserved in GEN1, FEN1, Rad2 and EXO1 (root mean square deviations of 0.9–1.1 Å for 161 Cα atoms, respectively). The residues corresponding to the tip of the thumb (residues 79–92), which are disordered in the GEN1 structure, likely form helix α4 upon DNA binding to the downstream interface as seen in human FEN1 and EXO1 (***Orans et al., 2011***; ***Tsutakawa et al., 2011***). The missing residues in GEN1 have 35.7% identity and 78.6% similarity (BLOSUM62 matrix) to the corresponding residues in FEN1 (90–103), which form helix α4 in the FEN1-DNA complex (compare ***Figure 2***). The same region is disordered in FEN1 when no DNA is bound (***Sakurai et al., 2005***). This indicates that also GEN1 undergoes such a disorder-to-order transition to form an arch with helices α4 and α6 upon substrate binding (***Patel et al., 2012***) and similar to the arrangement in T5 FEN (***Ceska et al., 1996***).

The activity of GEN1 depends on correct DNA positioning

GEN1 has versatile substrate recognition features, ranging from gaps, flaps, replication fork intermediates to HJs (***Ip et al., 2008***; ***Ishikawa et al., 2004***; ***Kanai et al., 2007***). To understand the functional relevance of the GEN1 structure for DNA recognition we performed a series of mutagenesis studies with single point mutations and truncated protein variants (***Figure 5*** and ***Figure 5—figure supplement 1/2***) to investigate the effect on the active site (D30N), upstream DNA binding (R54E), downstream DNA binding (C36E), arch at the downstream interface (R89E, R93E, H109E, F110E), and chromodomain (Δchromo, K404E, R406E). We performed nuclease assays by titrating different amounts of GEN1 to a fixed DNA concentration of 40 nM for 15 min and DNA cleavage products were analyzed by native electrophoresis (***Figure 5A*** and ***Figure 5—figure supplement 1/2***). We used an immobile HJ and a 5′ flap substrate side-by-side to facilitate the comparison of the effects on separate GEN1 functions. Notably, stoichiometric amounts of GEN1 were required to cleave HJ substrates whereas 5′ flaps were readily processed with catalytic amounts (***Figure 5A***).

The active site modification D30N showed that the cleavage activity on both HJ and 5′ flap substrates was lost in agreement with previously published data (***Ip et al., 2008***). According to our

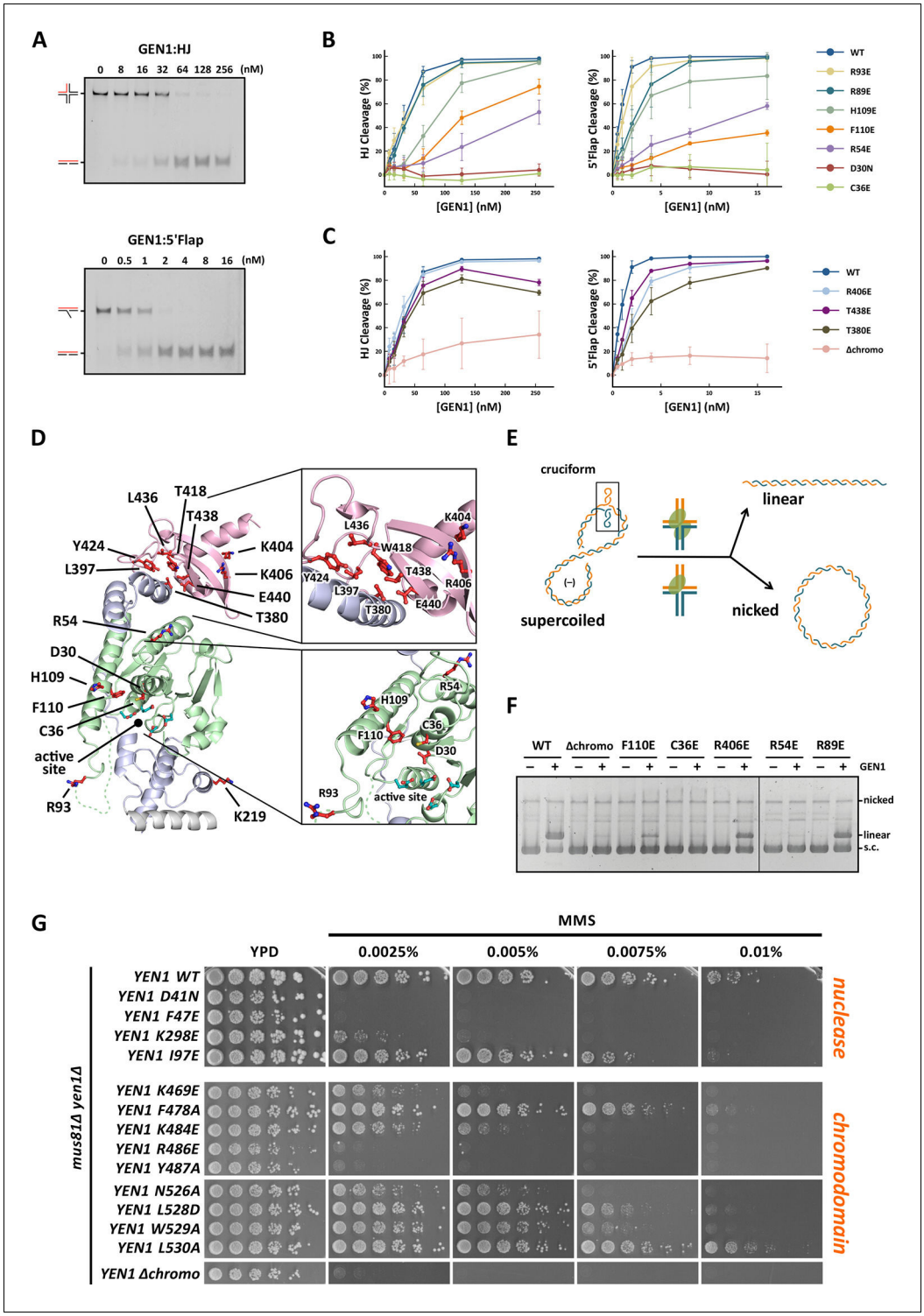


Figure 5. Functional analysis of GEN1. (**A**) Nuclease activity of GEN1 with HJ and 5'flap DNA. 40 nM 5' 6FAM-labeled substrates were mixed with indicated amounts of GEN1. Reactions were carried out at 37°C for 15 min, products were separated by native PAGE and analyzed with a phosphoimager. ***Figure 5—source data 1*** gives the sequences of DNA oligos used in biochemical assays and ***Figure 5—source data 3*** shows activity measurements. (**B**) Quantification of nuclease assays of wild type GEN1 and variants with mutated residues located at the protein-DNA interfaces. Percentage of cleavage was plotted against the enzyme concentration. Error bars depict the standard deviation calculated from at least three independent experiments. ***Figure 5—figure supplement 1*** shows representative gels from the PAGE analysis. (**C**) Quantification of nuclease assays of wild type GEN1 and

Figure 5 continued on next page

Figure 5 continued

variants with mutated residues located at the chromodomain. Error bars depict the standard deviation calculated from at least three independent experiments. ***Figure 5—figure supplement 2*** shows representative gels from the PAGE analysis. (**D**) GEN1 mutations used in this study. Locations of human GEN1 mutations used in biochemical assays and corresponding residues in yeast MMS survival assays are highlighted in red. Active site residues E134, E136, D155, D157 are marked in turquoise. (**E**) Schematic of the cruciform plasmid cleavage assay. A cruciform structure can be formed in plasmid pIRbke8^{mut}, which harbors an inverted-repeat sequence and is stabilized by negative supercoiling. Introducing two cuts across the junction point within the lifetime of the resolvase-junction complex yields linear products whereas sequential cleavage generates nicked products and the relaxed plasmid cannot be a substrate for the next cleavage. (**F**) Cruciform plasmid cleavage assay with different GEN1 variants. Plasmid pIRbke8^{mut} was treated with 256 nM GEN1 each and reactions were carried out at 37°C for 15 min. Supercoiled, linear and nicked plasmids were separated by native agarose gel electrophoresis and visualized with SYBR safe under UV light. (**G**) MMS survival assays with yeast *yen1* variants. The survival of *yen1* mutants was tested under a *yen1*Δ *mus81*Δ background with indicated amounts of MMS. The top part shows mutations at GEN1-DNA interfaces and the bottom part mutations at the chromodomain (compare ***Figure 5—figure supplement 3*** for all controls and expression tests). ***Figure 5—source data 2*** gives a list of all yeast strains.

DOI: 10.7554/eLife.12256.013

The following source data and figure supplements are available for figure 5:

Source data 1. Oligonucleotides used in biochemical assays.

DOI: 10.7554/eLife.12256.014

Source data 2. Yeast strains used for MMS survival assays.

DOI: 10.7554/eLife.12256.015

Source data 3. In vitro activity measurements of different GEN1$^{2\text{-}505}$ variants.

DOI: 10.7554/eLife.12256.016

Figure supplement 1. DNA cleavage assays of different GEN1 mutations.

DOI: 10.7554/eLife.12256.017

Figure supplement 2. DNA cleavage assays of different GEN1 fragments.

DOI: 10.7554/eLife.12256.018

Figure supplement 3. MMS survival assays with yeast *yen1* mutants.

DOI: 10.7554/eLife.12256.019

structure, R54 in helix α3 at the upstream interface fixes the substrate position by reaching into the minor DNA groove and we observed that R54E had a strongly reduced cleavage activity (~50%; ***Figure 5B***), indicating a key role in substrate positioning.

Residue C36 in helix α2 points towards the downstream interface and likely contacts the DNA upon binding (compare ***Figure 5D***). The corresponding FEN1 Y40, is a key residue stacking with the -1 base of the 5′ flap at the FEN1 active site (***Tsutakawa et al., 2011***). Therefore, we tested the cleavage ability of a GEN1^{C36E} and found that the mutant protein had completely lost its enzymatic activity for both, HJ and 5′ flap cleavage, to the same degree as the active site modification D30N (***Figure 5B***). This effect is stronger than for FEN1^{Y40A}, which showed only a partial loss in activity (***Tsutakawa et al., 2011***). Our results suggest that C36 provides a polar interface for orienting and guiding the cleaved strand towards the active site and the lower gateway.

We further tested a glutamate modification of the superfamily-conserved R89 and R93 located in the disordered part continuing to helix α6, presumably forming an arch (see above). The arch was shown to facilitate cleavage by clamping flap substrates in FEN1 and the modification R100A showed a strong decrease in the cleavage activity (***Patel et al., 2012***). The GEN1 R89E mutation, corresponding to residue R100 in FEN1, showed that the activity of GEN1 with a HJ substrate was not altered. In the case of a 5′ flap substrate, cleavage was slightly reduced and it reached to the full level at enzyme concentrations higher than 10 nM. The effect of the R93E modification was even less pronounced compared to R89E. In contrast, the cleavage of both 5′ flap and HJ substrates depended strongly on F110 at helix α6 (thumb), which points towards the active site. An F110E modification showed a reduction in cleavage by 25% for HJ substrates, and the effect was even stronger for 5′ flap substrates, where the activity is reduced by 65%. The equivalent position in FEN1 is V133 showing a critical involvement in stabilizing 5′ flap DNA by orienting the -1 nucleotide for catalysis (***Tsutakawa et al., 2011***). We have also tested the effect of modifying H109, which neighbors the critical F110. Even though it points away from the active site, a glutamate at this

position reduced 5′ flap cleavage to 83% and HJ cleavage recovered only at high substrate concentrations of 256 nM. Overall, the results suggest that F110 has a key position for DNA recognition and processing.

Coordinated cleavage of HJs

Classical HJ resolvases introduce two symmetrical incisions across the junction point by coordinating the action of two active sites. The first nick is rate-limiting and the second one takes place near-simultaneously and within the lifetime of the resolvase-DNA complex. This mechanism has been well studied for bacterial and bacteriophage HJ resolvases (***Fogg and Lilley, 2000***; ***Giraud-Panis and Lilley, 1997***; ***Pottmeyer and Kemper, 1992***; ***Shah et al., 1997***). Hence, it is thought that also GEN1 dimerizes upon binding to HJ substrates as indicated by coordinated cleavage and by an increase in hydrodynamic radius compared to protein alone (***Chan and West, 2015***; ***Rass et al., 2010***). In order to further examine the effect of GEN1 modifications on HJ cleavage, we used a cruciform plasmid cleavage assay to evaluate GEN1's nicking function, as illustrated in ***Figure 5E***. Here, the plasmid pIRbke8^{mut} served as a substrate that contains an inverted-repeat sequence extruding a cruciform structure when supercoiled (***Chan and West, 2015***; ***Lilley, 1985***; ***Rass et al., 2010***). Coordinated dual incision of the cruciform (by a dimer) leads to linear duplex products with slow migration, whereas uncoordinated cleavage (by monomeric enzymes) results in nicked plasmids that migrate even slower (***Figure 5F***). Cruciform structures are reabsorbed when the superhelical stress is released upon single nicking and the DNA cannot serve as a substrate anymore.

We observed that wild type GEN1 resolved cruciform structures into linear products (***Figure 5F***) in agreement with previous reports (***Chan and West, 2015***; ***Rass et al., 2010***). GEN1^{C36E} (downstream interface) and GEN1^{R54E} (upstream interface) showed only residual activity confirming their importance for HJ cleavage. The cruciform cleavage by F110E (thumb) was strongly reduced in line with our nuclease assays using small DNA substrates (***Figure 5B***). GEN1^{R89E} (disordered part of the arch) did not show any appreciable effect, which suggests that this part of the arch is not directly involved in HJ recognition. Taken together, our results suggest that the positioning of HJ junction substrates both at the upper and the lower gateway is critical for productive cleavage. Furthermore, none of the tested modifications at the different DNA interaction interfaces was able to uncouple the coordinated HJ cleavage.

The chromodomain of GEN1 facilitates efficient substrate cleavage

Agreeing with the structural significance for DNA binding, the truncation of the chromodomain (Δchromo, residues 2-389) showed a severe reduction (~3-fold) in HJ cleavage activity whereas all longer GEN1 fragments containing the chromodomain (2-464, 2-505 and 2-551) showed full activity (***Figure 5—figure supplement 2***). Interestingly, the effect of the chromodomain truncation is even more pronounced for 5′ flap DNA cleavage than for HJs, showing a 7-fold reduction compared to wild type (***Figure 5C***). The activity of GEN1 in the plasmid-based cruciform cleavage assay was also severely hampered in the absence of the chromodomain (***Figure 5F***) showing only a weak band for linear products and no increase for nicked plasmid, emphasizing the importance of the chromodomain for GEN1 activity.

Further, to test the influence of the positively charged side chains K404 and R406 on DNA binding, we introduced charge-reversal mutations to glutamates and assessed their nuclease activities. Even though K404 and R406 are within hydrogen-bonding distance to the DNA, K404E, and R406E showed no appreciable influence on GEN1's nuclease activity. Only a slight reduction in cleavage of 5′ flap substrates was observed for GEN1^{R406E}, whereas the processing of HJ substrates was not altered significantly (***Figure 5C***). This reinforces the conclusion from our structural observations that the chromodomain and the DNA interact through their backbones via van-der-Waals interactions.

Influence of phosphorylation-mimicking chromodomain modifications

PhosphoSitePlus (***Hornbeck et al., 2014***) lists two phosphorylation sites at residues T380 and T438 in GEN1 that were found in a T-cell leukemia and a glioblastoma cell line. These residues are located in helix α15 and at the rim of the aromatic cage, respectively. Both phosphorylation sites are positioned to interrupt hydrophobic interactions between helix α15 and the chromodomain (***Figure 5D*** and ***Figure 4F***). Therefore, we tested if the phosphorylation-mimicking modifications T380E and

T438E had an effect on GEN1's activity. At low enzyme concentrations (<50 nM) HJ cleavage was similar to that of wild-type protein but at high concentrations the activity declined to less than 80% (***Figure 5C***). For a 5' flap substrate, the assay showed consistently lower activity than wild type, recovering to about 80% cleavage at the highest enzyme concentration (***Figure 5C***). These results suggest that phosphorylation of GEN1 chromodomain residues may regulate DNA recognition and cleavage.

Physiological relevance of GEN1 interactions

To test the physiological relevance of the identified GEN1-DNA interactions, we investigated the survival of *Saccharomyces cerevisiae* mutant strains expressing variants of Yen1 (GEN1 homolog) after treatment with the DNA-damaging agent MMS (***Figure 5G*** and ***Figure 5—figure supplement 3/source data 2***). All Yen1 variants were expressed to a similar degree as endogenous Yen1, which was confirmed by Western Blot analysis (***Figure 5—figure supplement 3***). Because of the functional overlap of Mus81 and Yen1 in HR (***Blanco et al., 2010***) a double knockout (*yen1*Δ *mus81*Δ) was used and complemented with different variants of *Yen1*.

The control strain, complemented with wild type Yen1, survived MMS concentrations of up to 0.01%, consistent with the described hypersensitivity of *mus81*Δ mutants (***Blanco et al., 2010***; ***Interthal and Heyer, 2000***). In stark contrast, cells containing either the active site mutant Yen1-D41N (corresponding to $GEN1^{D30N}$) or the downstream interface mutant Yen1-F47E (corresponding to $GEN1^{C36E}$) did not grow even at an MMS concentration as low as 0.0025% (***Figure 5G***). After the expression of the upstream interface mutant Yen1-I97E (corresponding to $GEN1^{R54E}$) cells showed a slight but significant growth defect at high MMS concentrations (see panels for 0.0075% and 0.01% MMS in ***Figure 5G***). These results are therefore consistent with the in vitro cleavage results carried out with GEN1 mutants and showing a reduction in activity for R54E and no activity for C36E (see ***Figure 5C***). As a last mutant in the nuclease core, we tested the K298E mutation which is located in helix α10 of the H2TH motif in the downstream DNA-binding interface, and for which we were unable to obtain the corresponding $GEN1^{K219E}$ modification for cleavage assays (compare ***Figure 5D***). This mutant displayed a strong sensitivity towards MMS but lower than the one observed for the catalytic mutant, indicating that the mutant was partially functional in yeast (***Figure 5G***).

We next investigated the effect of mutations in the aromatic cage of Yen1's chromodomain (compare ***Figure 3***) and found that their severity was strongly position dependent. Mutation of R486E and Y487A in Yen1, both of which are located near the base of the cage, corresponding to the W418 position in GEN1 (see ***Figure 3C***), showed a strong effect on MMS sensitivity (see ***Figure 5G***), similar to the one observed for the catalytic mutant, presumably due to a dysfunctional chromodomain. In contrast, mutations located further outside of the core (F478A and K484E) led to a less pronounced MMS sensitivity. The same was true for the K469E variant, which corresponds to position R406 at the chromodomain-DNA interface in GEN1 (see ***Figure 3A*** and ***5F***), and for residues at the rim of the chromodomain (*yen1-N526A, yen1-L528D* and *yen1-W529A*), consistent with our in vitro observation for $GEN1^{T438E}$ (slightly reduced activity, ***Figure 5C***). No effect on MMS sensitivity was detected for *yen1-L530A*, which corresponds to a conserved glutamate in chromodomains (E440 in GEN1). Lastly, we found that the deletion of the chromodomain (Yen1-Δ452–560) lead to a severe phenotype comparable to the active site mutant Yen1-D41N (***Figure 5G*** and ***Figure 5—source data 2***). The Yen1 variant lacking the chromodomain was expressed to levels similar to the full-length protein and we therefore conclude that the chromodomain is crucial for the function of Yen1. Taken together, the functional data of Yen1 mutants in vivo and GEN1 mutants in vitro point towards an essential and evolutionary conserved role of the chromodomain in GEN1/Yen1 proteins.

Discussion

Implications of the chromodomain

The structure of the human GEN1 catalytic core provides the missing structural information in the Rad2/XPG family. The GEN1 structure complements recent reports on the structures of Rad2, EXO1 and FEN1, (***Miętus et al., 2014***; ***Orans et al., 2011***; ***Tsutakawa et al., 2011***). Thereby, it gives insights how relatively conserved nuclease domains recognize diverse substrates in a structure-

selective manner and act in different DNA maintenance pathways. In comparison with other Rad2/XPG nucleases, GEN1 shows many modifications on common structural themes that give the ability to recognize a diverse set of substrates including replication fork intermediates and HJs. The upstream DNA interface of GEN1 lacks the 'acid block' found in FEN1, instead it has a prominent groove at the same position (compare ***Figure 1***, 'upper gate') with a strategically positioned R54 nearby. Furthermore, the helical arch in GEN1 misses helix α5, which forms a cap structure in FEN1 and EXO1 that stabilizes 5' overhangs for cleavage. These features have implications for the recognition and cleavage of HJ substrates (see below). The most striking difference to other Rad2/XPG family members is that the GEN1 nuclease core is extended by a chromodomain, which provides an additional DNA anchoring point for the upstream DNA-binding interface. The evolutionarily conserved chromodomain is important for efficient substrate cleavage as we showed using truncation and mutation analyses. This finding opens new perspectives for the regulation of GEN1 and for its interactions with other proteins. Chromodomains serve as chromatin-targeting modules (reviewed in ***Blus et al., 2011***; ***Eissenberg, 2012***; ***Yap and Zhou, 2011***), general protein interaction elements (***Smothers and Henikoff, 2000***) as well as dimerization sites (***Canzio et al., 2011***; ***Cowieson et al., 2000***; ***Li et al., 2011***). These possibilities are particularly interesting, as chromatin targeting of proteins via chromodomains has been implicated in the DNA damage response. The chromatin remodeler CHD4 is recruited in response to DNA damage to decondense chromatin (reviewed in ***O'Shaughnessy and Hendrich, 2013***; ***Stanley et al., 2013***). The chromodomains in CHD4 distinguish the histone modifications H3K9me3 and H3K9ac and determine the way how downstream DSB repair takes place (***Ayrapetov et al., 2014***; ***Price and D'Andrea, 2013***). It is plausible that GEN1 uses its chromodomain not only as a structural module to securely bind DNA but also for targeting or regulatory purposes. Even though it was not possible to find any binding partner with a series of tested histone tail peptides, we cannot exclude that the chromodomain is used as an interaction motif or chromatin reader. It will therefore be interesting to extend our interaction analysis to a larger number of peptides and proteins. Interestingly, the modifications GEN1^{L397E} and GEN1^{Y424A} at the rim of the chromodomain did not alter DNA cleavage activity (***Figure 5—figure supplement 1***), however, mutations of residues at the rim of Yen1's chromodomain show a phenotype, suggesting an additional role like binding to an endogenous factor.

Another intriguing aspect of the chromodomain is that the conserved T438 at the rim of the aromatic cage and T380 at the closing helix α15 are both part of a casein kinase II consensus sequence for phosphorylation (Ser/Thr-X-X-Asp/Glu). ***Ayoub et al., 2008*** showed that the analogous threonine in the chromodomain of CBX1 is phosphorylated in response to DNA damage and phosphorylation disrupts the binding to H3K9me. We observed a reduction in DNA cleavage activity for the phosphorylation mimicking mutations T380E and T438E, which may suggest a regulatory role. They might function together and in combination with other modifications to provide a way of functional switching at the chromodomain. Furthermore, ***Blanco et al., 2014*** and ***Eissler et al., 2014*** recently identified several CDK phosphorylation sites in an insertion in the Yen1 chromodomain which affects HJ cleavage and together with phosphorylation of a nuclear localization signal (NLS) in the regulatory domain restricts Yen1's activity to anaphase. The insertion is not found in other chromodomains and it is extended in Yen1 compared to GEN1, which is lacking these phosphorylation sites (compare ***Figure 3A/B***). Notably, the activity of Yen1 is negatively regulated by CDK-dependent phosphorylation (***Blanco et al., 2014***; ***Chan and West, 2014***; ***Eissler et al., 2014***; ***Matos et al., 2011***), suggesting that the chromodomain is targeted by cell cycle kinases. It also provides a likely explanation for the different regulatory mechanisms found in GEN1 and Yen1 (***Blanco and Matos, 2015***; ***Chan and West, 2014***; ***Matos and West, 2014***). Exploration of the regulatory function of the GEN1 chromodomain will be an important topic to follow up, and this may lead to the understanding of the precise regulation mechanism of GEN1 as well as its substrate recognition under physiological conditions.

It is noteworthy that our analysis also revealed that the human transcription modulator AEBP2, which is associated with the polycomb repression complex 2 (PRC2), contains a chromo-barrel domain, which, to our knowledge, has not been reported so far.

Recognition of DNA substrates

The GEN1-DNA structure showed a considerable similarity to the other members of the Rad2/XPG family, and this facilitated the generation of a combined model to understand substrate recognition

of GEN1 (***Figure 6***). This was done by superimposing the protein part of the FEN1-DNA complex (PDB 3q8k) onto our GEN1 structure and extending the DNA accordingly (***Figure 6A/B***). Remarkably, the superimposition of the proteins aligns the DNA from the FEN1 structure in the same register as the DNA in the GEN1 complex at the upstream interface (***Figure 6A and 6B*** insert). Furthermore, the free 5′ and 3′ ends of the double flap DNA from the FEN1 structure point towards the lower and the upper gateway in GEN1, respectively (***Figure 6B***). We extended the GEN1 structure by homology modeling of the disordered residues 79-92 (helix α4) in GEN1 (***Figure 6B***). In addition to the similarity of this part to FEN1, the model readily showed the arrangement forming an arch structure. This would explain why GEN1 recognizes 5′ flap substrates efficiently, analogous to FEN1, as the arch can clamp a single-stranded DNA overhang for productive cleavage. This also explains why the F110E modification in the arch at helix α6 hampered 5′ flap cleavage severely. The

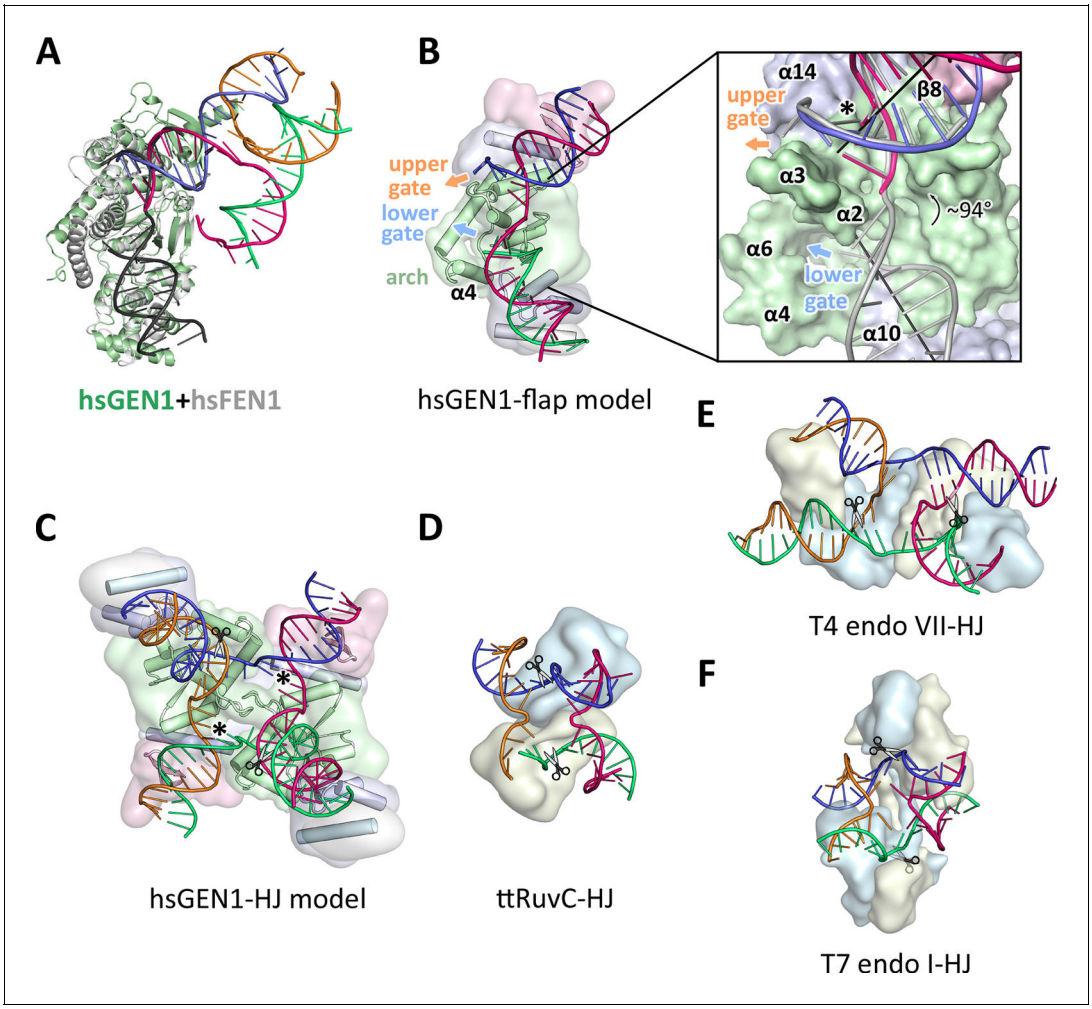


Figure 6. Substrate recognition features of GEN1. (**A**) Superposition of the protein part of the FEN1-DNA complex (PDB 3q8k, protein in gray, DNA in black) onto the GEN1-HJ complex (protein in green and the DNA strands in different colors). The FEN1-DNA aligns with the same register as the GEN1-DNA at the upstream interface. (**B**) Model for the recognition of a 5′ flap substrate by GEN1. The DNA was extended using the superimposition from **A**. Homology modeling suggests an additional helix α4 (disordered residues 79–92) forming an arch with helix α6. The protein is shown in a simplified surface representation with the same colors as in ***Figure 1*** and structural elements are highlighted. The insert shows a zoomed in view of the hydrophobic wedge with the modeled FEN1-DNA in gray. (**C**) Model for the dimerization of GEN1 upon binding to a HJ substrate based on the 5′ flap model in B. The monomers interlock via both arches (α4-α6) and the hydrophobic wedges (α2-α3) contact each other. (**D**) Structure of the *Thermus thermophilus* RuvC-HJ complex (PDB 4ld0). (**E**) Structure of the T4 endonuclease VII-HJ complex (PDB 2qnc). (**F**) Structure of the T7 endonuclease I-HJ complex (PDB 2pfj). Individual monomers are in surface representation, colored in light blue and beige, respectively. DNA strands are shown as ladders in different colors.

DOI: 10.7554/eLife.12256.020

side chain points directly towards the active site and likely disturbs the stabilization of a 5′ overhang for catalysis by charge repulsion. However, there are two features in GEN1 that vary from the arrangement in FEN1 and EXO1 considerably. Helix α6 is longer (24 instead of 15 residues) and helix α5 is missing in GEN1. As a result the arch points away from the DNA rather than forming a 'cap' structure as it is observed in FEN1 and EXO1 (***Orans et al., 2011***; ***Tsutakawa et al., 2011***). Furthermore, the modified arch in GEN1 provides an opening, marked as 'lower gate' in ***Figure 6B***. These differences are likely the basis for GEN1's versatile DNA recognition features.

Implications of an adjustable hatch in GEN1 for substrate discrimination

The diverging orientation of the arch (helices α4 and α6) in GEN1 compared to the one in FEN1 and EXO1 (helices α4, α5, and α6) may have thus significance for the recognition of HJ substrates. By pointing away from the active site the arch provides an opening to accommodate unpaired, single-stranded DNA to pass along the arch at the lower gate (groove between α2 and α4) (***Figure 6B*** 'lower gate') from one GEN1 monomer to the upper gate (groove between α2-α3 and α14) (***Figure 6B*** 'upper gate') of the other within a GEN1 dimer (***Figure 6B/C***). R54 is perfectly positioned at the minor groove to guide the second cleavage strand to pass through the upper gate (compare ***Figure 4*** and ***Figure 6B/C***, marked with a asterisk). In FEN1, this position is occupied by the 'acid block', which stabilizes a single 3′ flap of the unpaired substrate (***Tsutakawa et al., 2011***) and it would not accommodate longer 3′ DNA overhangs. In our model, two GEN1 monomers come together crosswise upon HJ binding (***Figure 6C***). The helical arches of both proteins likely provide additional protein-protein interactions as well as protein-DNA contacts by packing against the backbone of opposite DNA arms (***Figure 6C***). As a result, the GEN1 dimer orients both active sites symmetrically across the junction point resembling the situation in bacterial RuvC (***Figure 6D***; ***Bennett and West, 1995a***; ***Górecka et al., 2013***). This arrangement would ensure that both incisions are introduced within the lifetime of the GEN1-HJ complex as observed biochemically by us and others (***Chan and West, 2015***; ***Rass et al., 2010***). The mechanism likely works in a coordinated nick-and-counter-nick fashion, as shown for bacterial or bacteriophage HJ resolvases (***Fogg and Lilley, 2000***; ***Giraud-Panis and Lilley, 1997***; ***Pottmeyer and Kemper, 1992***; ***Shah et al., 1997***) and recently for GEN1 (***Chan and West, 2015***).

The distance between both gates is bridged by unpaired bases in our GEN1-HJ model. This view is supported by the observation that FEN1 unpairs two bases near the active site through interactions with the hydrophobic wedge leading to strongly bent DNA arms between the upstream and downstream DNA interfaces. This mechanism seems to be a common feature of Rad2/XPG nucleases (***Finger et al., 2013***; ***Grasby et al., 2012***; ***Tsutakawa et al., 2011***). Consistent with this view, the bacterial RuvC resolvase (***Figure 6D***) has also been shown to unfold HJ junctions (***Bennett and West, 1995b***; ***Górecka et al., 2013***). In the case of GEN1, the critical step would be the assembly of the dimer around the junction point in a highly restraint way and the introduction of the first nick. This releases the tension on the complex like a spring leading to an immediate second cut and subsequent disassembly of the GEN1-HJ complex. Furthermore, a HJ does not provide free DNA ends and adopts a structure that intrinsically restrains its degrees of freedom, thus inhibiting cleavage by a single GEN1 monomer. Altogether we speculate that the arch (helix α4-α6) acts like a lever or hatch switching between flap and HJ recognition modes. When a free 5′ end is available it closes and clamps the flap, thus positions the DNA for cleavage. For the case of a HJ substrate, the arch adopts an open conformation, allowing unpaired, single-stranded DNA to pass, while preventing the correct positioning of the DNA for catalysis at first. HJ cleavage is inhibited until a second GEN1 monomer binds. This mechanism differs from the one used by bacterial or bacteriophage HJ resolvases, which act as obligate dimers binding to DNA substrates in a concerted way (compare ***Figure 6D–F***). Our model for DNA cleavage by GEN1 describes a conformational switch provided by a flexible arch that can discriminate between substrates containing free 5′ ends or those with a restraint structure like HJs. This aspect may explain our observation that GEN1 cleaves 5′ flap DNA catalytically while stoichiometric amounts are required for HJ substrates (***Figure 5A–C***). Using a switchable hatch in a spring-loaded mechanism would be an efficient way of preventing a single cut at a HJ junction while allowing GEN1 to adapt to recognize various DNA substrates and perform different functional roles. Thus, GEN1 may have an intrinsic safety mechanism that ensures symmetrical dual incision across a branch point. Further studies have to address the exact engagement mechanism.

GEN1 in a biological context

GEN1's biological role is not fully understood yet. Yeast cells are viable without the GEN1 homolog Yen1 even in the presence of DNA damaging agents as the Mus81-Eme1 complex can complement the defect (compare ***Figure 5—figure supplement 3***; ***Blanco et al., 2010***). Consistently, both proteins can cleave 5′ flaps and HJ substrates in vitro. However, GEN1 can cleave intact HJs symmetrically whereas MUS81-EME1 is much more efficient with nicked DNA four-way junctions (***Castor et al., 2013***; ***Wyatt et al., 2013***). ***Matos et al., 2011*** suggested that Yen1/GEN1 might serve as a backup enzyme to resolve persistent HJs that have eluded other mechanisms of joint molecule removal before cytokinesis.

Our analysis infers that HJ cleavage is slower than 5′ flap cleavage (***Figure 5B/C***), bringing interesting implications for a safety control of GEN1's activity. GEN1 may have to assemble in an accurate way before it can cleave a HJ. Likewise, it increases GEN1's persistence time on HJs and opens a window for branch migration for extending the length of recombined stretches of DNA. Moreover, GEN1 recognizes various DNA substrates, which may point towards a general role in processing substrates in different DNA maintenance pathways. GEN1 has been shown to cleave replication fork intermediates, and it is implicated in the resolution of replication-induced HJs (***Garner et al., 2013***; ***Sarbajna et al., 2014***). Like MUS81-EME1, it might also be important for the processing of fragile sites to ensure proper chromosome segregation (***Ying et al., 2013***). These functions have to be tested systematically to understand GEN1's biological role. In this context, the regulation of GEN1 is an important factor and needs to be explored. Our study identified a chromodomain extending the GEN1 nuclease core that might have a role in regulating the enzyme. An open question is the function and architecture of the remaining 444 amino acids at the C-terminus of GEN1. They are thought to regulate the nuclease activity and control subcellular localization (***Blanco et al., 2014***; ***Chan and West, 2014***; ***García-Luis et al., 2014***). It is very likely that new interaction sites and post-translational modifications in this region will be discovered in future. The presented structure together with additional studies will help to unravel these questions and to obtain a comprehensive view of the functions of the Rad2/XPG nucleases.

Materials and methods

Experimental procedures

Protein expression and purification

Wild type human GEN1 and truncations thereof (residues 2-551, 2-505, 2-464, 2-389) were amplified by PCR from IMAGE clone 40125755 (Mammalian Gene collection, natural variant S92T, S310N, UniProtID Q17RS7) and cloned into a self-made ligation-independent cloning vector with various C-terminal tags followed by His8. Truncated versions were designed based on limited proteolysis in combination with domain prediction and functional assays to determine the smallest yet active fragment. The N-terminal methionine was cleaved by cellular methionyl-aminopeptidase, which is an essential requirement in the Rad2/XPG family as the N-terminus (conserved residue G2) folds towards the active site. Mutations were introduced by site-directed mutagenesis using Phusion Polymerase (NEB, Frankfurt/Main, Germany). All recombinant proteins were expressed in the *E. coli* BL21(DE3) pRIL strain (MerckMillipore, Darmstadt, Germany). Cells were grown at 37°C until mid-log phase and induced overnight with 0.2 mM IPTG at 16°C. Cells were harvested by centrifugation and resuspended in lysis buffer containing 1x phosphate buffered saline (PBS) with additional 500 mM NaCl, 10% (v/v) glycerol, 2 mM DTT, 1 mM EDTA, 1 µM leupeptin, 1 µM pepstatin A, 0.1 mM AEBSF and 2 µM aprotinin and lyzed by sonication. Cell debris was removed by centrifugation (75 600 *g* for 45 min), the clarified lysate was applied onto Complete HisTag Nickel resin (Roche Diagnostics, Mannheim, Germany) and washed with buffer A consisting of 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 10% (v/v) glycerol, 2 mM DTT and followed by a chaperone wash step with 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 2 mM ATP, 5 mM $MgCl_2$, 10% (v/v) glycerol and 2 mM DTT. The protein was eluted with buffer A containing 300 mM imidazole. The tag was cleaved, followed by cation exchange chromatography using a HiTrap SP HP column (GE Healthcare, Freiburg, Germany) with a linear gradient from 150 mM to 450 mM NaCl. Peak fractions were pooled and further purified by size-exclusion chromatography on a HiLoad 16/60 Superdex 200 (GE Healthcare) equilibrated with 20 mM Tris-HCl

pH 7.5, 100 mM NaCl, 5%(v/v) glycerol, 0.1 mM EDTA and 2 mM TCEP. Peak fractions were pooled, concentrated, flash-frozen in liquid nitrogen and stored at -80˚C.

Crystallization and data collection

GEN1$^{2\text{-}505\ D30N}$ and DNA (4w1010-1 GAATTCCGGATTAGGGATGC, 4w1010-2 GCATCCCTAAGC TCCATCGT, 4w1010-3 ACGATGGAGCCGCTAGGCTC, 4w1010-4 GAGCCTAGCGTCCGGAATTC) were mixed at a molar ratio of 2:1.1 at a final protein concentration of 14 mg/ml including 1 mM $MgCl_2$ and co-crystallized by sitting drop vapor diffusion. Drops were set up by mixing sample with mother liquor consisting of 100 mM MES-NaOH pH 6.5 and 200 mM NaCl at a 2:1 ratio at room temperature. Crystals grew within 2 days, and several iterations of streak seeding were needed for obtaining diffraction quality crystals. For data collection, crystals were stepwise soaked in 10%, 20%, and 30% (v/v) glycerol in 100 mM MES-NaOH pH 6.5, 200 mM NaCl and 5% PEG 8000 and flash-frozen in liquid nitrogen. Diffraction data were collected at beamline PXII of the Swiss Light Source (SLS, Villigen, Switzerland) at 100 K with a Pilatus 6M detector. In order to obtain phase information, crystals were soaked for 10–30 min in 1 mM $[Ta_6Br_{12}]Br_2$, flash-frozen and data were collected at the Ta L(III)-edge. In addition, seleno-methionine (SeMet)-substituted protein was expressed in M9 media supplemented with SeMet, purified, and crystallized according to the protocol above and data were collected at the Se K-edge.

Structure determination and refinement

All data were processed with XDS (***Table 1***, ***Kabsch, 2010***). HKL2MAP (***Pape and Schneider, 2004***) found 12 tantalum and 8 selenium positions, which were used in a combined MIRAS strategy (multiple isomorphous replacement with anomalous scattering) in autoSHARP (***Vonrhein, et al., 2007***) to determine the structure of the GEN1-HJ complex. The obtained solvent-flattened experimental map was used to build a model with PHENIX (***Adams et al., 2010***) combined with manual building. The structure was then further refined by iterative rounds of manual building in COOT (***Emsley and Cowtan, 2004***), refinement with PHENIX and assisted by the PDB_REDO server (***Joosten, et al., 2014***). The structure was visualized and analyzed in PYMOL (***Delano, 2002***). Electrostatic surface potentials were calculated with PDB2PQR (***Dolinsky et al., 2004***) and APBS (***Baker et al., 2001***).

Nuclease assay

All DNA substrates (***Figure 5—source data 1***) were synthesized by Eurofins/MWG (Ebersberg, Germany), resuspended in annealing buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl, 0.1 mM EDTA), annealed by heating to 85˚C for 5 min and slow-cooling to room temperature. Different amounts of GEN1 proteins (as indicated) were mixed with 40 nM 6FAM-labeled DNA substrates in 20 mM Tris-HCl pH 8.0, 50 ng/µl bovine serum albumin (BSA) and 1 mM DTT. Reactions were initiated by adding 5 mM $MgCl_2$, incubated at 37˚C for 15 min and terminated by adding 15 mM EDTA, 0.3% SDS and further, DNA substrates were deproteinized using 1 mg/ml proteinase K at 37˚C for 15 min. Products were separated by 8% 1x TBE native polyacrylamide gel electrophoresis, the fluorescence signal detected with a Typhoon FLA 7000 phosphoimager (GE Healthcare), quantified with IMAGEJ (GE Healthcare) and visualized by GNUPLOT (***Williams et al., 2015***).

Cruciform plasmid cleavage assay

The cruciform plasmid pIRbke8^{mut} was a gift from Stephen West's lab (***Rass et al., 2010***), and it was originally prepared by David Lilley's lab (***Lilley, 1985***). 50 ng/µl plasmid were mixed with 20 mM Tris-HCl pH 8.0, 50 mM potassium glutamate, 5 mM $MgCl_2$, 50 ng/µl BSA and 1 mM DTT and pre-warmed at 37˚C for 1 hr to induce the formation of a cruciform structure. Reactions were initiated by adding indicated amounts of GEN1, incubated at 37˚C for 15 min and stopped as for DNA cleavage assays. The products were separated by 1% 1xTBE native agarose gel electrophoresis, stained with SYBR safe (Life Technologies, Darmstadt, Germany) and visualized under UV light.

Sequence alignments and phylogenetic analysis

Sequences of GEN1 proteins from different organisms as well as all human chromodomain proteins were aligned to the human GEN1 sequence using the programs HHPRED (***Söding et al., 2005***), PSI-BLAST and further by manual adjustments. Alignments were tested by back-searches against RefSeq

or HMM databases. A phylogenetic tree was calculated by the program PHYML with 100 bootstraps using the alignment in ***Figure 3—figure supplement 1*** and a BLOSUM62 substitution model. The tree was displayed with DENDROSCOPE (***Huson and Scornavacca, 2012***).

Histone peptide pull-down assay

The GEN1 chromodomain with a C-terminal His8-tag was immobilized on complete HisTag Nickel resin and washed twice with binding buffer consisting of 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 5% glycerol, 0.1 mM EDTA, 0.05% (v/v) Tween-20 and 2 mM TCEP. Peptide mixtures containing 0.4 µM fluorescein labeled histone peptides were incubated with beads at 4˚C for 1 hr and washed twice with binding buffer. Immobilized proteins were eluted with binding buffer supplemented with 300 mM imidazole and separated on 20% SDS-PAGE. Fluorescein-labeled peptides were visualized by detecting the fluorescence signal with a Typhoon FLA 7000 phosphoimager (GE Healthcare).

Yeast genetics and MMS survival assay in *Saccharomyces cerevisiae*

All yeast strains are based on W303 Rad5+ (see ***Figure 5—source data 2*** for a complete list). *yen1*△ or *yen1*△ *mus81*△ strains were transformed with an integrative plasmid expressing mutant versions of *YEN1*. Freshly grown over-night cultures were diluted to $1x10^7$ cells/ml. 5-fold serial dilutions were spotted on YPD plates with/without MMS (methyl methanesulphonate, concentrations as indicated) and incubated for 2 days at 30˚C. The expression of 3FLAG-tagged Yen1 constructs was verified by SDS-PAGE and Western Blot analysis. Proteins were detected using a mouse monoclonal anti-FLAG M2-peroxidase (HRP) antibody (Sigma-Aldrich, München, Germany).

Database entry

The coordinates of the human GEN1-Holliday junction complex have been deposited in the Protein Data Bank (PDB code 5t9j).

Acknowledgements

We would like to thank Naoko Mizuno, Michael J. Taschner and Esben Lorentzen for many scientific discussions and critical reading of the manuscript. We are grateful for the support by the Department of Structural Cell Biology at the MPI of Biochemistry (MPIB), particularly for the help with screening by the Crystallization Facility and Claire Basquin for assistance with biophysical analysis. The Microchemistry Core Facility of the MPIB provided mass spectrometry analysis. We would like to thank Jürg Müller for providing histone tail peptides for binding tests and Stephen C. West for providing the plasmid pIRbke8^{mut} for functional assays. Furthermore, we would like to acknowledge the professional assistance of the beamline staff at the Swiss Light Source, Villigen, Switzerland, during data collection. The Max Planck Society for the Advancement of Science supported this research through the Otto Hahn Program (BP) and the Max Planck Research Group Leader Program (CB).

Additional information

Funding

Funder	Grant reference number	Author
Max-Planck-Gesellschaft	Max Planck Research Group Leader Program	Shun-Hsiao Lee Maren Felizitas Klügel Christian Biertümpfel
Max-Planck-Gesellschaft	Otto Hahn Program	Lissa Nicola Princz Boris Pfander
Max-Planck-Gesellschaft	MPI of Biochemistry	Bianca Habermann

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

S-HL, CB, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article, Contributed unpublished essential data or reagents; LNP, MFK, Acquisition of data, Contributed unpublished essential data or reagents; BH, Acquisition of data, Analysis and interpretation of data; BP, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article

Additional files

Major datasets

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database, license, and accessibility information
Lee S-H, Biertumpfel C	2016	Crystal Structure of human GEN1 in complex with Holliday junction DNA in the upper interface	http://www.rcsb.org/pdb/explore/explore.do?structureId=5T9J	Publicly available at the RCSB Protein Data Bank (accession no. 5T9J)

References

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography* **66**:213–221. doi: 10.1107/S0907444909052925

Andersen SL, Kuo HK, Savukoski D, Brodsky MH, Sekelsky J, Lichten M. 2011. Three structure-selective endonucleases are essential in the absence of BLM helicase in Drosophila. *PLoS Genetics* **7**:e1002315. doi: 10.1371/journal.pgen.1002315

Ayoub N, Jeyasekharan AD, Bernal JA, Venkitaraman AR. 2008. HP1-beta mobilization promotes chromatin changes that initiate the DNA damage response. *Nature* **453**:682–686. doi: 10.1038/nature06875

Ayrapetov MK, Gursoy-Yuzugullu O, Xu C, Xu Y, Price BD. 2014. DNA double-strand breaks promote methylation of histone H3 on lysine 9 and transient formation of repressive chromatin. *Proceedings of the National Academy of Sciences* **111**:9169–9174. doi: 10.1073/pnas.1403565111

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences* **98**:10037–10041. doi: 10.1073/pnas.181342398

Bennett RJ, West SC. 1995a. RuvC protein resolves Holliday junctions via cleavage of the continuous (noncrossover) strands. *Proceedings of the National Academy of Sciences* **92**:5635–5639. doi: 10.1073/pnas.92.12.5635

Bennett RJ, West SC. 1995b. Structural analysis of the RuvC-Holliday junction complex reveals an unfolded junction. *Journal of Molecular Biology* **252**:213–226. doi: 10.1006/jmbi.1995.0489

Benson FE, West SC. 1994. Substrate specificity of the Escherichia coli RuvC protein. Resolution of three- and four-stranded recombination intermediates. *The Journal of Biological Chemistry* **269**:5195–5201.

Biertümpfel C, Yang W, Suck D. 2007. Crystal structure of T4 endonuclease VII resolving a Holliday junction. *Nature* **449**:616–620. doi: 10.1038/nature06152

Blanco MG, Matos J, Rass U, Ip SC, West SC. 2010. Functional overlap between the structure-specific nucleases Yen1 and Mus81-Mms4 for DNA-damage repair in S. cerevisiae. *DNA Repair* **9**:394–402. doi: 10.1016/j.dnarep.2009.12.017

Blanco MG, Matos J, West SC. 2014. Dual control of Yen1 nuclease activity and cellular localization by Cdk and Cdc14 prevents genome instability. *Molecular Cell* **54**:94–106. doi: 10.1016/j.molcel.2014.02.011

Blanco MG, Matos J. 2015. Hold your horSSEs: controlling structure-selective endonucleases MUS81 and Yen1/GEN1. *Frontiers in Genetics* **6**:1–11. doi: 10.3389/fgene.2015.00253

Blus BJ, Wiggins K, Khorasanizadeh S. 2011. Epigenetic virtues of chromodomains. *Critical Reviews in Biochemistry and Molecular Biology* **46**:507–526. doi: 10.3109/10409238.2011.619164

Canzio D, Chang EY, Shankar S, Kuchenbecker KM, Simon MD, Madhani HD, Narlikar GJ, Al-Sady B. 2011. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Molecular Cell* **41**:67–81. doi: 10.1016/j.molcel.2010.12.016

Castor D, Nair N, Déclais AC, Lachaud C, Toth R, Macartney TJ, Lilley DM, Arthur JS, Rouse J. 2013. Cooperative control of holliday junction resolution and DNA repair by the SLX1 and MUS81-EME1 nucleases. *Molecular Cell* **52**:221–233. doi: 10.1016/j.molcel.2013.08.036

Cejka P, Plank JL, Bachrati CZ, Hickson ID, Kowalczykowski SC. 2010. Rmi1 stimulates decatenation of double Holliday junctions during dissolution by Sgs1-Top3. *Nature Structural & Molecular Biology* **17**:1377–1382. doi: 10.1038/nsmb.1919

Cejka P, Plank JL, Dombrowski CC, Kowalczykowski SC. 2012. Decatenation of DNA by the S. cerevisiae Sgs1-Top3-Rmi1 and RPA complex: a mechanism for disentangling chromosomes. *Molecular Cell* **47**:886–896. doi: 10.1016/j.molcel.2012.06.032

Ceska TA, Sayers JR, Stier G, Suck D. 1996. A helical arch allowing single-stranded DNA to thread through T5 5'-exonuclease. *Nature* **382**:90–93. doi: 10.1038/382090a0

Chan YW, West S. 2015. GEN1 promotes Holliday junction resolution by a coordinated nick and counter-nick mechanism. *Nucleic Acids Research* **43**:10882–10892. doi: 10.1093/nar/gkv1207

Chan YW, West SC. 2014. Spatial control of the GEN1 Holliday junction resolvase ensures genome stability. *Nature Communications* **5**:4844. doi: 10.1038/ncomms5844

Cowieson NP, Partridge JF, Allshire RC, McLaughlin PJ. 2000. Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Current Biology* **10**:517–525. doi: 10.1016/S0960-9822(00)00467-X

Delano WL. *The PyMOL molecular graphics system*. 2002.

Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. 2004. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Research* **32**:W665–W667. doi: 10.1093/nar/gkh381

Eissenberg JC. 2012. Structural biology of the chromodomain: form and function. *Gene* **496**:69–78. doi: 10.1016/j.gene.2012.01.003

Eissler CL, Mazón G, Powers BL, Savinov SN, Symington LS, Hall MC. 2014. The Cdk/cDc14 module controls activation of the Yen1 holliday junction resolvase to promote genome stability. *Molecular Cell* **54**:80–93. doi: 10.1016/j.molcel.2014.02.012

Emsley P, Cowtan K. 2004. *Coot*: model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography* **60**:2126–2132. doi: 10.1107/S0907444904019158

Fekairi S, Scaglione S, Chahwan C, Taylor ER, Tissier A, Coulon S, Dong MQ, Ruse C, Yates JR, Russell P, Fuchs RP, McGowan CH, Gaillard PH. 2009. Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell* **138**:78–89. doi: 10.1016/j.cell.2009.06.029

Finger LD, Patel N, Beddows A, Ma L, Exell JC, Jardine E, Jones AC, Grasby JA. 2013. Observation of unpaired substrate DNA in the flap endonuclease-1 active site. *Nucleic Acids Research* **41**:9839–9847. doi: 10.1093/nar/gkt737

Fogg JM, Lilley DM. 2000. Ensuring productive resolution by the junction-resolving enzyme RuvC: large enhancement of the second-strand cleavage rate. *Biochemistry* **39**:16125–16134. doi: 10.1021/bi001886m

García-Luis J, Clemente-Blanco A, Aragón L, Machín F. 2014. Cdc14 targets the Holliday junction resolvase Yen1 to the nucleus in early anaphase. *Cell Cycle* **13**:1392–1399. doi: 10.4161/cc.28370

Garner E, Kim Y, Lach FP, Kottemann MC, Smogorzewska A. 2013. Human GEN1 and the SLX4-associated nucleases MUS81 and SLX1 are essential for the resolution of replication-induced Holliday junctions. *Cell Reports* **5**:207–215. doi: 10.1016/j.celrep.2013.08.041

Giraud-Panis MJ, Lilley DM. 1997. Near-simultaneous DNA cleavage by the subunits of the junction-resolving enzyme T4 endonuclease VII. *The EMBO Journal* **16**:2528–2534. doi: 10.1093/emboj/16.9.2528

Górecka KM, Komorowska W, Nowotny M. 2013. Crystal structure of RuvC resolvase in complex with Holliday junction substrate. *Nucleic Acids Research* **41**:9945–9955. doi: 10.1093/nar/gkt769

Grasby JA, Finger LD, Tsutakawa SE, Atack JM, Tainer JA. 2012. Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases. *Trends in Biochemical Sciences* **37**:74–84. doi: 10.1016/j.tibs.2011.10.003

Hadden JM, Déclais AC, Carr SB, Lilley DM, Phillips SE. 2007. The structural basis of Holliday junction resolution by T7 endonuclease I. *Nature* **449**:621–624. doi: 10.1038/nature06158

Heyer WD. 2015. Regulation of recombination and genomic maintenance. *Cold Spring Harbor Perspectives in Biology* **7**:a016501. doi: 10.1101/cshperspect.a016501

Holliday R. 1964. A mechanism for gene conversion in fungi. *Genetical Research* **5**:282–304. doi: 10.1017/S0016672300001233

Holm L, Rosenström P. 2010. Dali server: conservation mapping in 3D. *Nucleic Acids Research* **38**:W545–W549. doi: 10.1093/nar/gkq366

Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. 2015. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Research* **43**:D512–D520. doi: 10.1093/nar/gku1267

Huson DH, Scornavacca C. 2012. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Systematic Biology* **61**:1061–1067. doi: 10.1093/sysbio/sys062

Interthal H, Heyer WD. 2000. MUS81 encodes a novel helix-hairpin-helix protein involved in the response to UV- and methylation-induced DNA damage in Saccharomyces cerevisiae. *Molecular and General Genetics MGG* **263**:812–827. doi: 10.1007/s004380000241

Ip SC, Rass U, Blanco MG, Flynn HR, Skehel JM, West SC. 2008. Identification of Holliday junction resolvases from humans and yeast. *Nature* **456**:357–361. doi: 10.1038/nature07470

Ira G, Malkova A, Liberi G, Foiani M, Haber JE. 2003. Srs2 and Sgs1-Top3 suppress crossovers during double-strand break repair in yeast. *Cell* **115**:401–411. doi: 10.1016/S0092-8674(03)00886-9

Ishikawa G, Kanai Y, Takata K, Takeuchi R, Shimanouchi K, Ruike T, Furukawa T, Kimura S, Sakaguchi K. 2004. DmGEN, a novel RAD2 family endo-exonuclease from Drosophila melanogaster. *Nucleic Acids Research* **32**:6251–6259. doi: 10.1093/nar/gkh962

Joosten RP, Long F, Murshudov GN, Perrakis A. 2014. The *PDB_REDO* server for macromolecular structure model optimization. *IUCrJ* **1**:213–220. doi: 10.1107/S2052252514009324

Kabsch W. 2010. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D Biological Crystallography* **66**:133–144. doi: 10.1107/S0907444909047374
Kanai Y, Ishikawa G, Takeuchi R, Ruike T, Nakamura R, Ihara A, Ohashi T, Takata K, Kimura S, Sakaguchi K, Nakamura Ryo-ichi, Takata Kei-ichi. 2007. DmGEN shows a flap endonuclease activity, cleaving the blocked-flap structure and model replication fork. *FEBS Journal* **274**:3914–3927. doi: 10.1111/j.1742-4658.2007.05924.x
Lee Bi BI, Nguyen LH, Barsky D, Fernandes M, Wilson DM. 2002. Molecular interactions of human Exo1 with DNA. *Nucleic Acids Research* **30**:942–949. doi: 10.1093/nar/30.4.942
Li J, Li Z, Ruan J, Xu C, Tong Y, Pan PW, Tempel W, Crombet L, Min J, Zang J, Gay N. 2011. Structural basis for specific binding of human MPP8 chromodomain to histone H3 methylated at lysine 9. *PLoS ONE* **6**:e25104. doi: 10.1371/journal.pone.0025104
Lieber MR. 1997. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *BioEssays* **19**:233–240. doi: 10.1002/bies.950190309
Lilley DM, White MF. 2001. The junction-resolving enzymes. *Nature Reviews Molecular Cell Biology* **2**:433–443. doi: 10.1038/35073057
Lilley DM. 1985. The kinetic properties of cruciform extrusion are determined by DNA base-sequence. *Nucleic Acids Research* **13**:1443–1465. doi: 10.1093/nar/13.5.1443
Matos J, Blanco MG, Maslen S, Skehel JM, West SC. 2011. Regulatory control of the resolution of DNA recombination intermediates during meiosis and mitosis. *Cell* **147**:158–172. doi: 10.1016/j.cell.2011.08.032
Matos J, West SC. 2014. Holliday junction resolution: regulation in space and time. *DNA Repair* **19**:176–181. doi: 10.1016/j.dnarep.2014.03.013
Mięt us M, Nowak E, Jaciuk M, Kustosz P, Studnicka J, Nowotny M. 2014. Crystal structure of the catalytic core of Rad2: insights into the mechanism of substrate binding. *Nucleic Acids Research* **42**:10762–10775. doi: 10.1093/nar/gku729
Muñoz IM, Hain K, Déclais AC, Gardiner M, Toh GW, Sanchez-Pulido L, Heuckmann JM, Toth R, Macartney T, Eppink B, Kanaar R, Ponting CP, Lilley DM, Rouse J. 2009. Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. *Molecular Cell* **35**:116–127. doi: 10.1016/j.molcel.2009.06.020
Nishino T, Ishino Y, Morikawa K. 2006. Structure-specific DNA nucleases: structural basis for 3D-scissors. *Current Opinion in Structural Biology* **16**:60–67. doi: 10.1016/j.sbi.2006.01.009
Orans J, McSweeney EA, Iyer RR, Hast MA, Hellinga HW, Modrich P, Beese LS. 2011. Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell* **145**:212–223. doi: 10.1016/j.cell.2011.03.005
O'Shaughnessy A, Hendrich B. 2013. CHD4 in the DNA-damage response and cell cycle progression: not so NuRDy now. *Biochemical Society Transactions* **41**:777–782. doi: 10.1042/BST20130027
Pape T, Schneider TR. 2004. *HKL2MAP* : a graphical user interface for macromolecular phasing with *SHELX* programs. *Journal of Applied Crystallography* **37**:843–844. doi: 10.1107/S0021889804018047
Patel N, Atack JM, Finger LD, Exell JC, Thompson P, Tsutakawa S, Tainer JA, Williams DM, Grasby JA. 2012. Flap endonucleases pass 5′-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5′-ends. *Nucleic Acids Research* **40**:4507–4519. doi: 10.1093/nar/gks051
Petronczki M, Siomos MF, Nasmyth K. 2003. Un ménage à quatre: the molecular biology of chromosome segregation in meiosis. *Cell* **112**:423–440 . doi: 10.1016/S0092-8674(03)00083-7
Pottmeyer S, Kemper B. 1992. T4 endonuclease VII resolves cruciform DNA with nick and counter-nick and its activity is directed by local nucleotide sequence. *Journal of Molecular Biology* **223**:607–615. doi: 10.1016/0022-2836(92)90977-R
Price BD, D'Andrea AD. 2013. Chromatin remodeling at DNA double-strand breaks. *Cell* **152**:1344–1354. doi: 10.1016/j.cell.2013.02.011
Putnam CD, Hayes TK, Kolodner RD. 2009. Specific pathways prevent duplication-mediated genome rearrangements. *Nature* **460**:984–989. doi: 10.1038/nature08217
Rass U, Compton SA, Matos J, Singleton MR, Ip SC, Blanco MG, Griffith JD, West SC. 2010. Mechanism of Holliday junction resolution by the human GEN1 protein. *Genes & Development* **24**:1559–1569. doi: 10.1101/gad.585310
Sakurai S, Kitano K, Yamaguchi H, Hamada K, Okada K, Fukuda K, Uchida M, Ohtsuka E, Morioka H, Hakoshima T. 2005. Structural basis for recruitment of human flap endonuclease 1 to PCNA. *The EMBO Journal* **24**:683–693. doi: 10.1038/sj.emboj.7600519
Sarbajna S, Davies D, West SC. 2014. Roles of SLX1-SLX4, MUS81-EME1, and GEN1 in avoiding genome instability and mitotic catastrophe. *Genes & Development* **28**:1124–1136. doi: 10.1101/gad.238303.114
Sarbajna S, West SC. 2014. Holliday junction processing enzymes as guardians of genome stability. *Trends in Biochemical Sciences* **39**:409–419. doi: 10.1016/j.tibs.2014.07.003
Schwacha A, Kleckner N. 1995. Identification of double Holliday junctions as intermediates in meiotic recombination. *Cell* **83**:783–791. doi: 10.1016/0092-8674(95)90191-4
Shah R, Cosstick R, West SC. 1997. The RuvC protein dimer resolves Holliday junctions by a dual incision mechanism that involves base-specific contacts. *The EMBO Journal* **16**:1464–1472. doi: 10.1093/emboj/16.6.1464
Smothers JF, Henikoff S. 2000. The HP1 chromo shadow domain binds a consensus peptide pentamer. *Current Biology* **10**:27–30. doi: 10.1016/S0960-9822(99)00260-2
Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* **33**:W244–W248. doi: 10.1093/nar/gki408

Stanley FKT, Moore S, Goodarzi AA. 2013. CHD chromatin remodelling enzymes and the DNA damage response. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **750**:31–44. doi: 10.1016/j.mrfmmm.2013.07.008

Svendsen JM, Harper JW. 2010. GEN1/Yen1 and the SLX4 complex: Solutions to the problem of Holliday junction resolution. *Genes & Development* **24**:521–536. doi: 10.1101/gad.1903510

Svendsen JM, Smogorzewska A, Sowa ME, O'Connell BC, Gygi SP, Elledge SJ, Harper JW. 2009. Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. *Cell* **138**:63–77. doi: 10.1016/j.cell.2009.06.030

Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. 1983. The double-strand-break repair model for recombination. *Cell* **33**:25–35. doi: 10.1016/0092-8674(83)90331-8

Tomlinson CG, Atack JM, Chapados B, Tainer JA, Grasby JA. 2010. Substrate recognition and catalysis by flap endonucleases and related enzymes. *Biochemical Society Transactions* **38**:433–437. doi: 10.1042/BST0380433

Tsutakawa SE, Classen S, Chapados BR, Arvai AS, Finger LD, Guenther G, Tomlinson CG, Thompson P, Sarker AH, Shen B, Cooper PK, Grasby JA, Tainer JA. 2011. Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* **145**:198–211. doi: 10.1016/j.cell.2011.03.004

Tsutakawa SE, Lafrance-Vanasse J, Tainer JA. 2014. The cutting edges in DNA repair, licensing, and fidelity: DNA and RNA repair nucleases sculpt DNA to measure twice, cut once. *DNA Repair* **19**:95–107. doi: 10.1016/j.dnarep.2014.03.022

Vonrhein C, Blanc E, Roversi P, Bricogne G. 2007. Automated structure solution with autoSHARP. *Methods in Molecular Biology* **364**:215–245. doi: 10.1385/1-59745-266-1:215

Wakasugi M, Reardon JT, Sancar A. 1997. The non-catalytic function of XPG protein during dual incision in human nucleotide excision repair. *Journal of Biological Chemistry* **272**:16030–16034. doi: 10.1074/jbc.272.25.16030

Wechsler T, Newman S, West SC. 2011. Aberrant chromosome morphology in human cells defective for Holliday junction resolution. *Nature* **471**:642–646. doi: 10.1038/nature09790

West SC, Blanco MG, Chan YW, Matos J, Sarbajna S, Wyatt HDM. 2015. Resolution of Recombination Intermediates: Mechanisms and Regulation. *Cold Spring Harbor Symposia on Quantitative Biology* **80**:103–109. doi: 10.1101/sqb.2015.80.027649

Williams T, Kelley C, Campbell J, Cunningham R, Denholm D, Elber G, Fearick R, Grammes C, Hart L, Hecking L. *Gnuplot 5.0.0*. 2015.

Wu L, Hickson ID. 2003. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426**:870–874. doi: 10.1038/nature02253

Wyatt HDM, Sarbajna S, Matos J, West SC. 2013. Coordinated Actions of SLX1-SLX4 and MUS81-EME1 for Holliday Junction Resolution in Human Cells. *Molecular Cell* **52**:234–247. doi: 10.1016/j.molcel.2013.08.035

Yang W. 2011. Nucleases: diversity of structure, function and mechanism. *Quarterly Reviews of Biophysics* **44**:1–93. doi: 10.1017/S0033583510000181

Yap KL, Zhou MM. 2011. Structure and mechanisms of lysine methylation recognition by the chromodomain in gene transcription. *Biochemistry* **50**:1966–1980. doi: 10.1021/bi101885m

Ying S, Minocherhomji S, Chan KL, Palmai-Pallag T, Chu WK, Wass T, Mankouri HW, Liu Y, Hickson ID. 2013. MUS81 promotes common fragile site expression. *Nature Cell Biology* **15**:1001–1007. doi: 10.1038/ncb2773

Figures and figure supplements

Human Holliday junction resolvase GEN1 uses a chromodomain for efficient DNA recognition and cleavage

Shun-Hsiao Lee *et al*

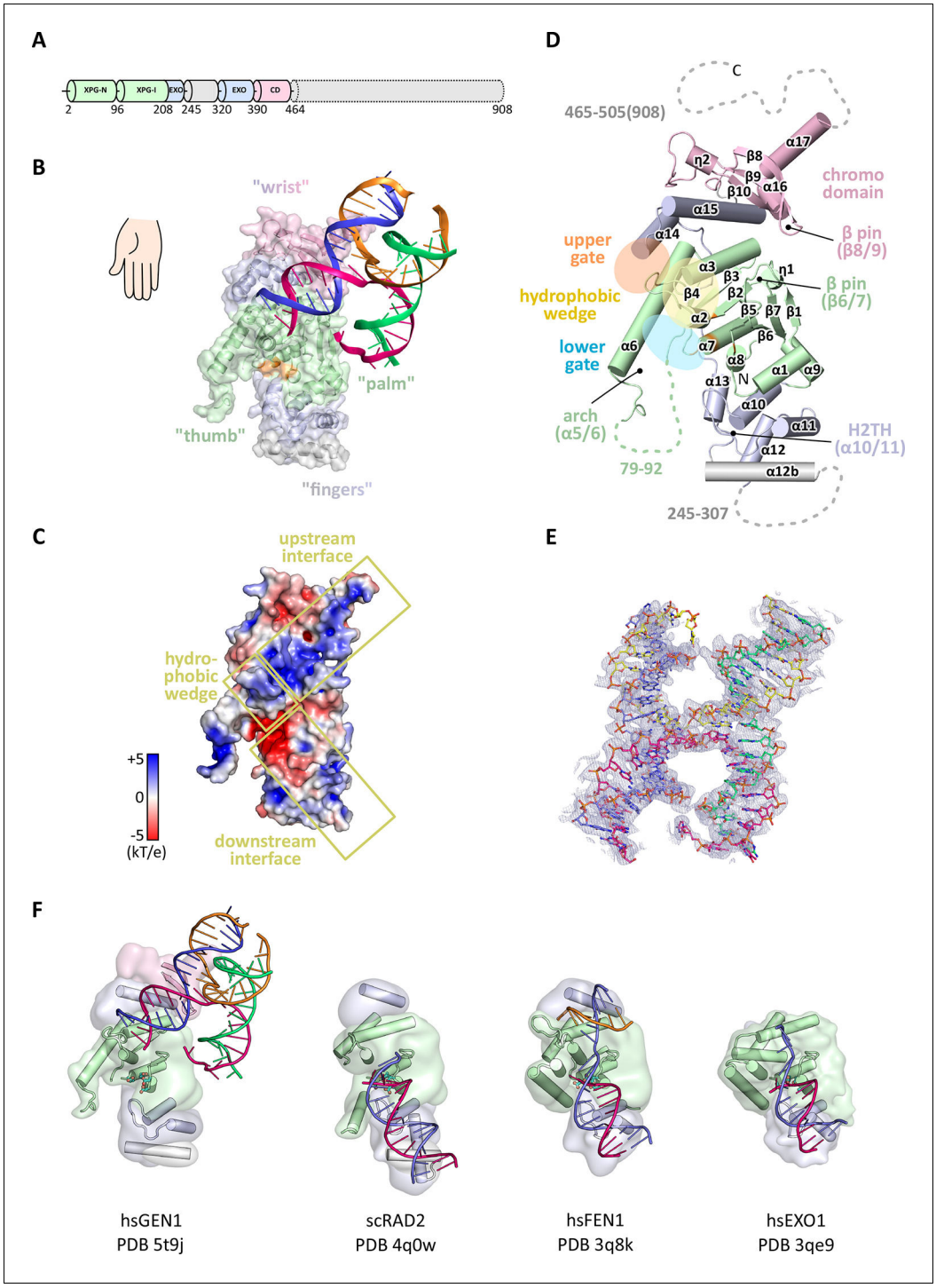


Figure 1. Architecture of human GEN1. (**A**) Domain architecture of human GEN1. The structurally unknown regulatory domain (residues 465–908) is shown with dotted lines. (**B**) Overview of the catalytic core of GEN1 in complex with HJ DNA. The protein resembles the shape of a downwards-pointing right hand with helix α6 as the thumb. The protein is depicted in half transparent surface representation with secondary structure elements underneath. The DNA is shown in ladder representation with individual strands in different colors. The coloring of GEN1 follows domain boundaries: intertwining XPG-N and XPG-I in green, 5′->3′ exonuclease C-terminal domain (EXO) in blue, chromodomain in pink, unassigned regions in gray. Active site residues (E134, E136, D155, D157) are highlighted in orange. (**C**) Electrostatic surface potential of GEN1. The coloring follows the potential from -5 (red) to +5 kT/e (blue). The DNA-binding interfaces and the position of the hydrophobic wedge are marked in

Figure 1 continued on next page

Figure 1 continued

yellow. (**D**) Secondary structure elements of the catalytic core of GEN1 in cartoon representation with the same colors as before. Dotted lines represent parts that are not resolved in the crystal structure. The numbering follows a unified scheme for the Rad2/XPG family (compare ***Figure 2***) for α-helices, β-sheets and 3_{10}-helices (η). (**E**) Experimental electron density map (autoSHARP, solvent flattened, contoured at 1σ) drawn around the HJ in the GEN1 complex. The DNA model is shown in ball-stick representation with carbon atoms of individual strands in different colors (yellow, light blue, magenta, green) and oxygen atoms in red, phosphor atoms in orange, nitrogen atoms in dark blue. (**F**) Structural comparison of Rad2/XPG family nucleases. Proteins are shown in a simplified surface representation with important structural elements in cartoon representation and DNA in ladder representation. The color scheme is the same as in **B**. ***Figure 1—figure supplement 1*** shows the content of the asymmetric unit.

DOI: 10.7554/eLife.12256.003

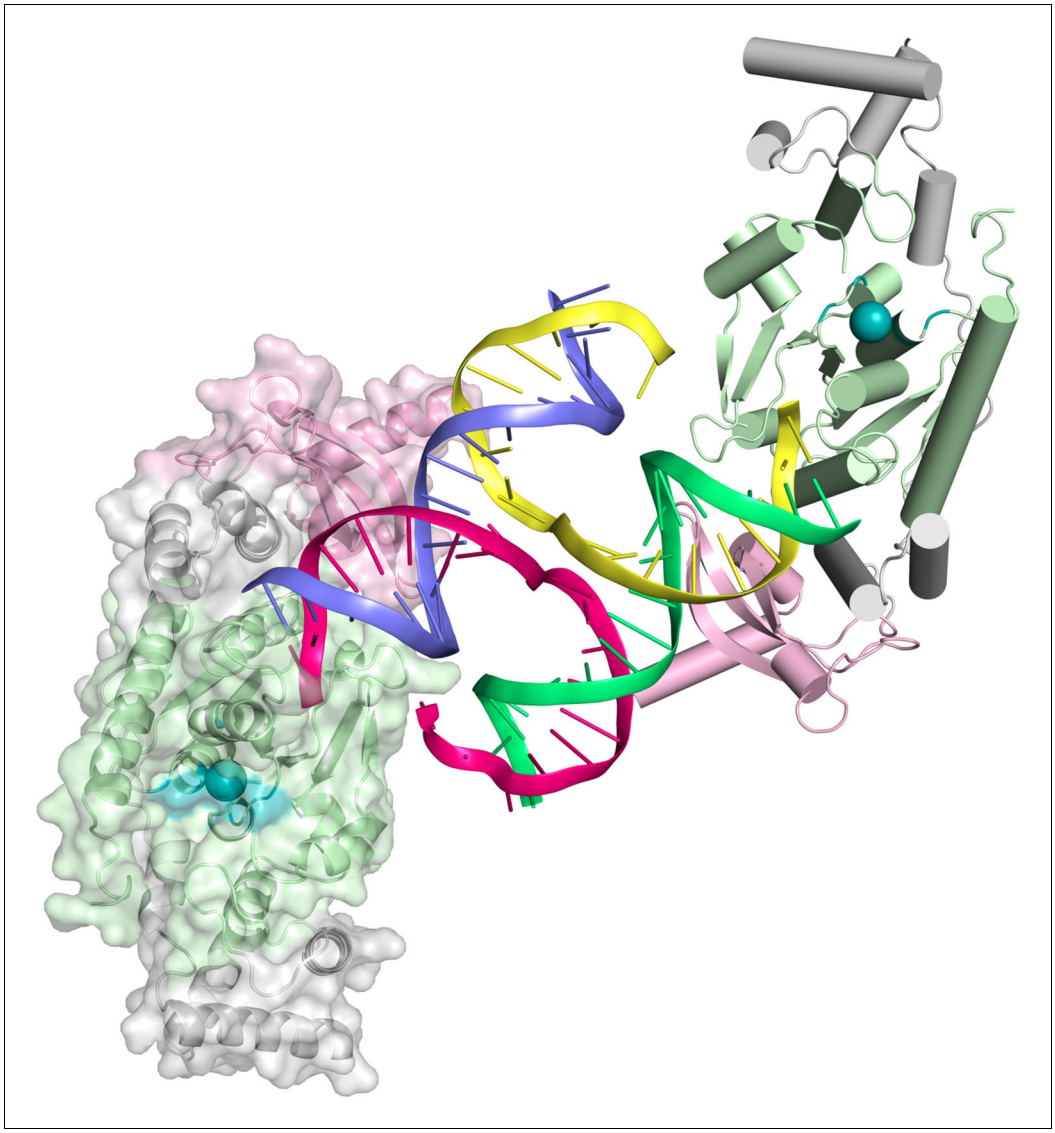


Figure 1—figure supplement 1. Content of the asymmetric unit of the GEN1-HJ crystal. One protein monomer is shown in surface representation with secondary structure cartoons underneath, the other one only in cartoon representation with α-helices as cylinders and β-strands as arrows. The HJ DNA bridges between two protein monomers in the asymmetric unit. The active sites are labeled with a turquoise ball each.
DOI: 10.7554/eLife.12256.004

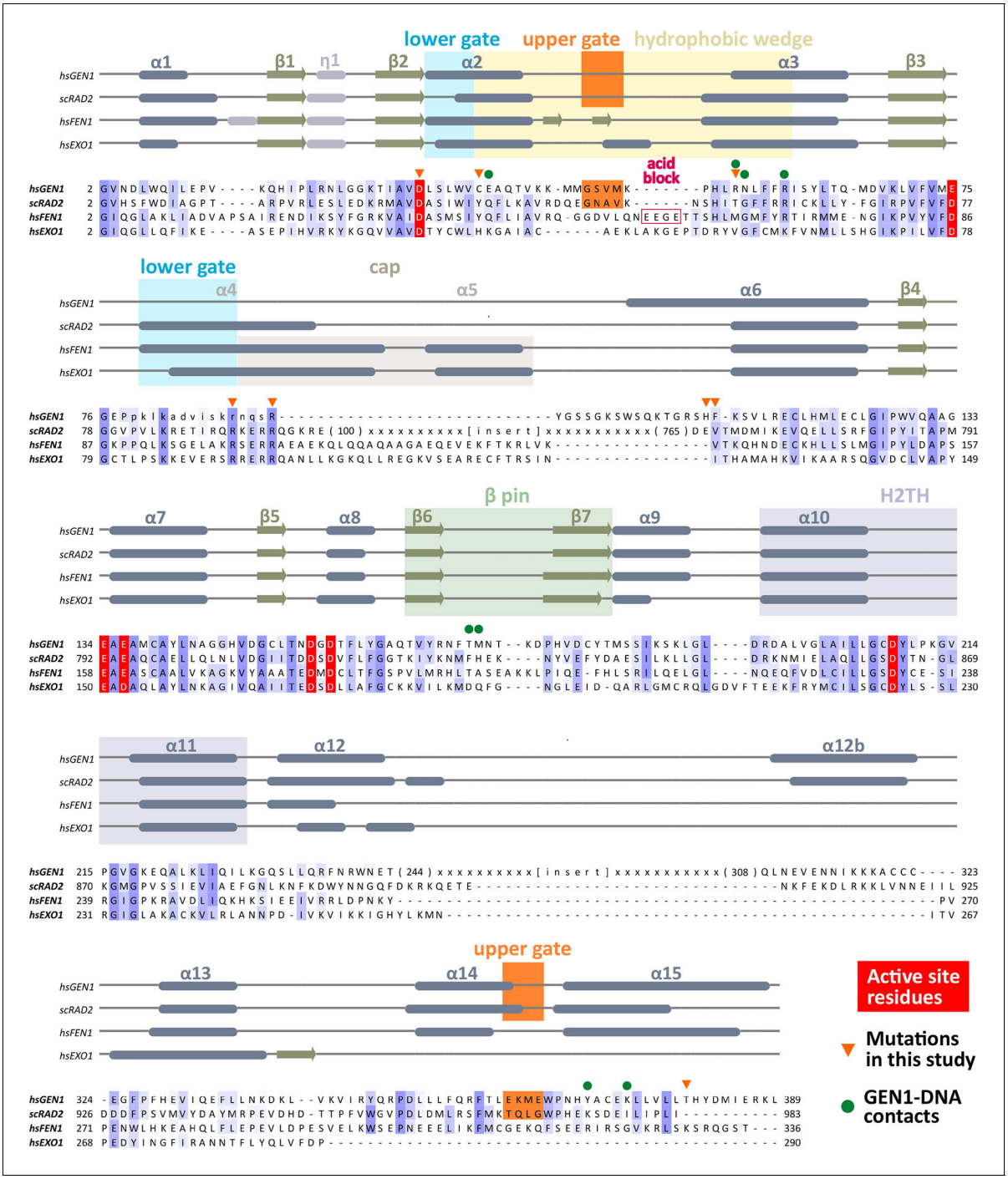


Figure 2. Alignment of the nuclease cores of Rad2/XPG-family proteins. The alignment is based on known crystal structures: human GEN1 (PDB 5t9j, this study), yeast Rad2 (PDB 4q0w), human FEN1 (PDB 3q8k), human EXO1 (3qe9). Secondary structure elements are depicted on top of the sequence with dark blue bars for α-helices, light blue bars for 3_{10}-helices and green arrows for β-sheets. The numbering follows a unified scheme for the superfamily. Functional elements are labeled and described in the main text. Sequences are colored by similarity (BLOSUM62 score) and active site residues are marked in red. Mutations analyzed in this study are marked with an orange triangle and DNA contacts found in the human GEN1–HJ structure have a dark green dot. Disordered or missing parts in the structures are labeled in small letters or with x.

DOI: 10.7554/eLife.12256.006

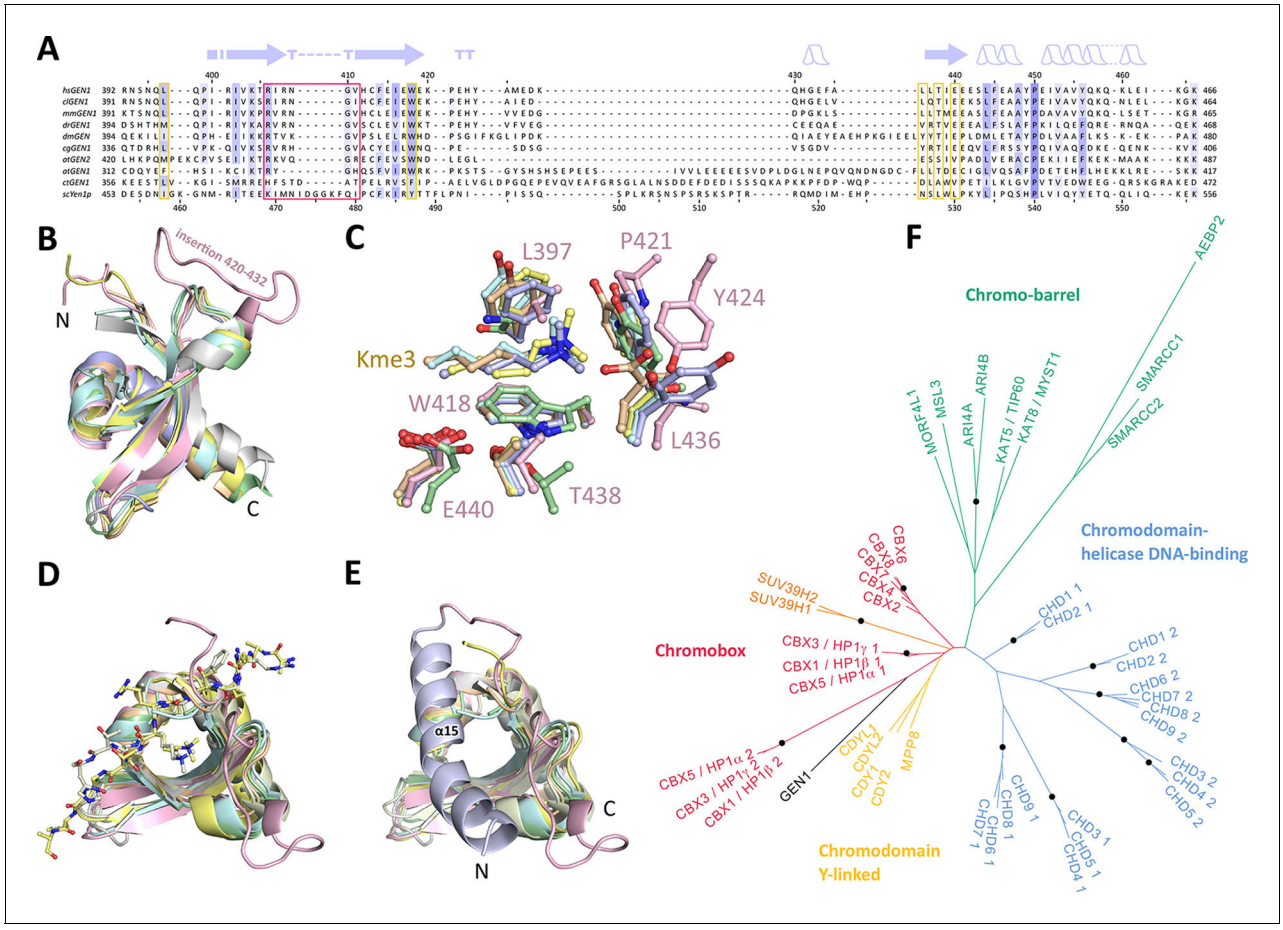


Figure 3. Chromodomain comparison. (**A**) Sequence alignment of GEN1 chromodomains from different organisms: hsGEN1 (*Homo sapiens*), clGEN1 (*Canis lupus*), mmGEN1 (*Mus musculus*), drGEN1 (*Danio rerio*), atGEN1/2 (*Arabidopsis thaliana*), cgGEN1 (*Crassostrea gigas*), scYEN1 (*Saccharomyces cerevisiae*). The presence of a chromodomain is conserved from yeast to human with *Caenorhabditis elegans* as an exception. Secondary structure elements of the GEN1 chromodomain are shown on top. The sequence coloring is based on a similarity matrix (BLOSUM62). The corresponding positions of the DNA-interaction site in human GEN1 is marked with a red box and residues of the aromatic cage are highlighted with a yellow box. (**B**) GEN1 has a canonical chromodomain fold of three antiparallel beta-sheets packed against an α-helix. (**C**) The arrangement of the aromatic cage in GEN1 is comparable to other chromodomains but less aromatic and slightly larger. (**D**) The superposition of different chromodomains places cognate binding peptides of hsMPP8 and mmCBX7 (and others) into the aromatic cage. (**E**) The aromatic cage of GEN1 is closed by helix α15. Panels **B–D** show the chromodomains of hsGEN1 (pink, PDB 5t9j), hsCBX3 (gray, PDB 3kup) hsSUV39H1 (green, PDB 3mts), hsMPP8 (yellow, PDB 3lwe), dmHP1a (orange, chromo shadow PDB 3p7j), dmRHINO (cyan, PDB 4quc/3r93), mmCBX7 (light blue, PDB 4x3s; compare ***Figure 3—source data 1***). (**F**) Phylogenetic tree of all known human chromodomains. GEN1 is distantly related to the CBX chromo-shadow domains and CDY chromodomains. The corresponding alignment for calculating the phylogenetic tree is shown in ***Figure 3—figure supplement 1***. GEN1 is colored in black, chromobox (CBX) proteins are colored in red, interspersed by SUV39H histone acetylases (orange) and chromodomain Y-linked (CDY) proteins (yellow). Chromo-barrel domain proteins are colored in green and chromodomain-helicase DNA-binding (CHD) proteins are in blue. Chromodomains and chromo-shadow domains from the same protein are labeled with 1 and 2, respectively. Stable branches with boostrap values equal or higher than 0.8 are marked with a black dot. The binding of the GEN1 chromodomain to a set of histone peptides was tested but no interaction was detected (***Figure 3—source data 2*** and ***Figure 3—figure supplement 2***).

DOI: 10.7554/eLife.12256.007

The following source data is available for figure 3:

Source data 1. Proteins found in a DALI search.

DOI: 10.7554/eLife.12256.008

Source data 2. N-terminally fluorescein-labeled peptides used for chromodomain binding assays.

DOI: 10.7554/eLife.12256.009

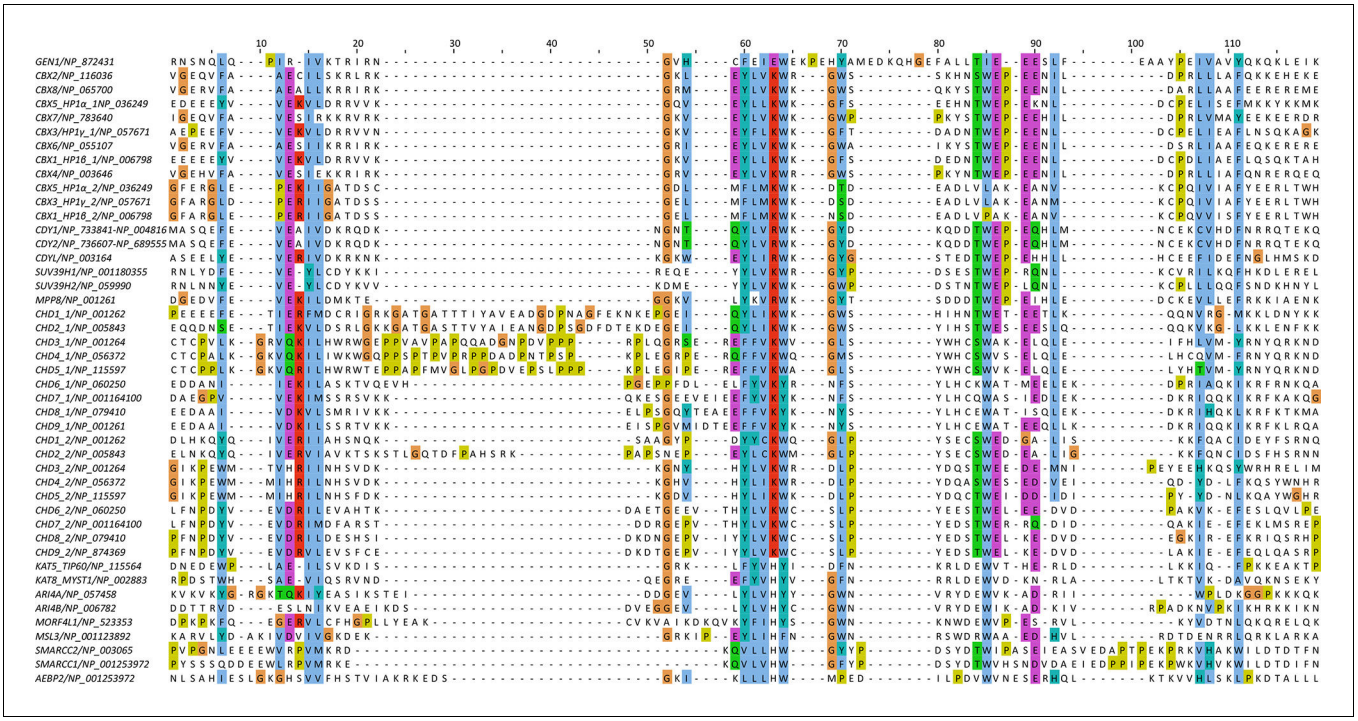


Figure 3—figure supplement 1. Sequence alignment of all known human chromodomains. The alignment was used to calculate the phylogenetic tree in ***Figure 3F***. Colors follow the CLUSTAL X coloring scheme.

DOI: 10.7554/eLife.12256.010

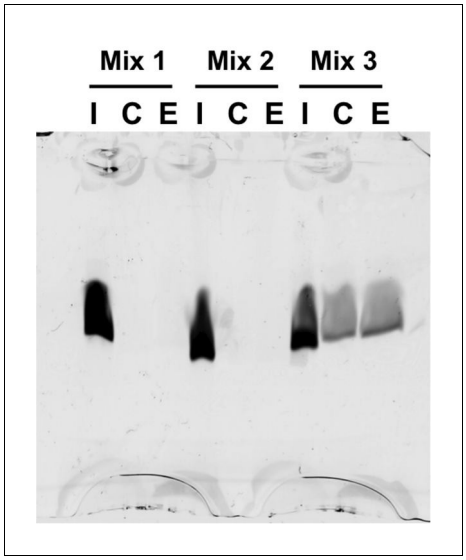


Figure 3—figure supplement 2. Histone peptide pull-down assay. Nickel resin-immobilized GEN1 chromodomain was incubated with the mixtures of fluorescein-labeled histone peptides, washed, bound peptides eluted and separated by 20% SDS-PAGE. Mix 1 and 2 did not show any binding, and non-specific binding to the resin was found with Mix 3. The smearing of the bands is due to the small size of the peptides (~1.5 kDa). I, C and E represent input, resin control and elution, respectively. Mix 1: H3K9, H3K9me1, H3K9me2, and H3K9me3. Mix 2: H3K27, H3K27me1, H3K27me2, and H3K27me3. Mix 3: H3K36me1, H3K36me2, H3K36me3, and H3K36Ac.
DOI: 10.7554/eLife.12256.011

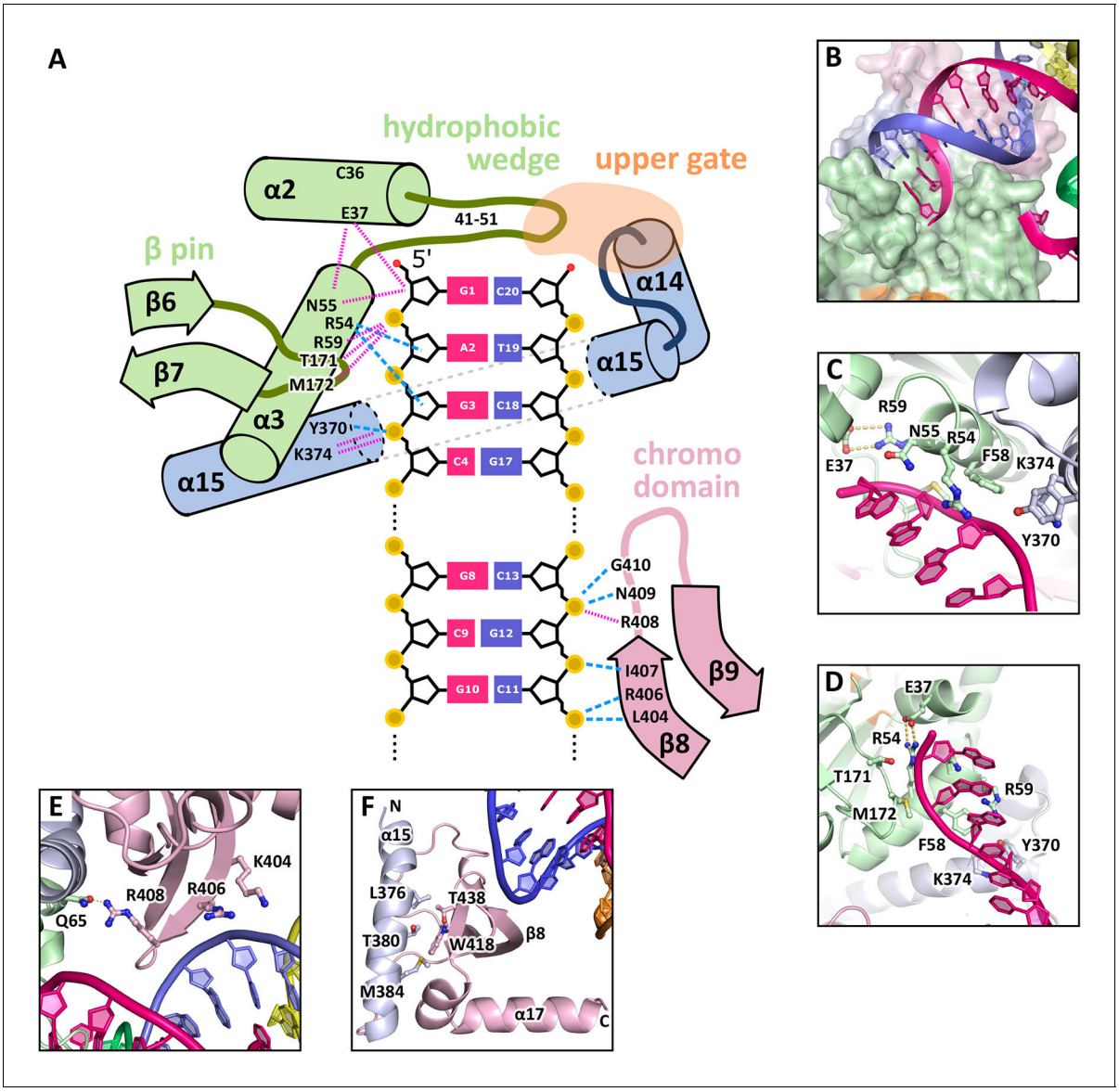


Figure 4. DNA interactions in the GEN1-DNA complex. (**A**) Schematic of the GEN1-DNA interactions at the upstream interface. The coloring is the same as in ***Figure 1***. The nuclease core (green and blue) interacts with the uncleaved strand and the chromodomain (pink) contacts the complementary strand. Hydrogen bonds are shown with blue dashed lines and van-der-Waals contacts are in red dotted lines. (**B**) Interactions at the hydrophobic wedge. The end of the DNA double helix docks onto the hydrophobic wedge formed by helices α2 and α3. (**C**/**D**) Interactions with the uncleaved strand in two views. All key residues form sequence-independent contacts to the DNA backbone. R54 reaches into the minor groove of the DNA. The complementary DNA strand has been removed for clarity (**E**/**F**) Interactions of the chromodomain with the complementary strand in two views. The backbone of residues 406–410 (β-hairpin β8-β9) abuts the DNA backbone. R406 has a supporting role in the interaction and R408 forms a polar interaction with Q65, which establishes a connection between the chromodomain and the nuclease core. Helix α15 makes hydrophobic interactions with the aromatic cage and thus blocks it.

DOI: 10.7554/eLife.12256.012

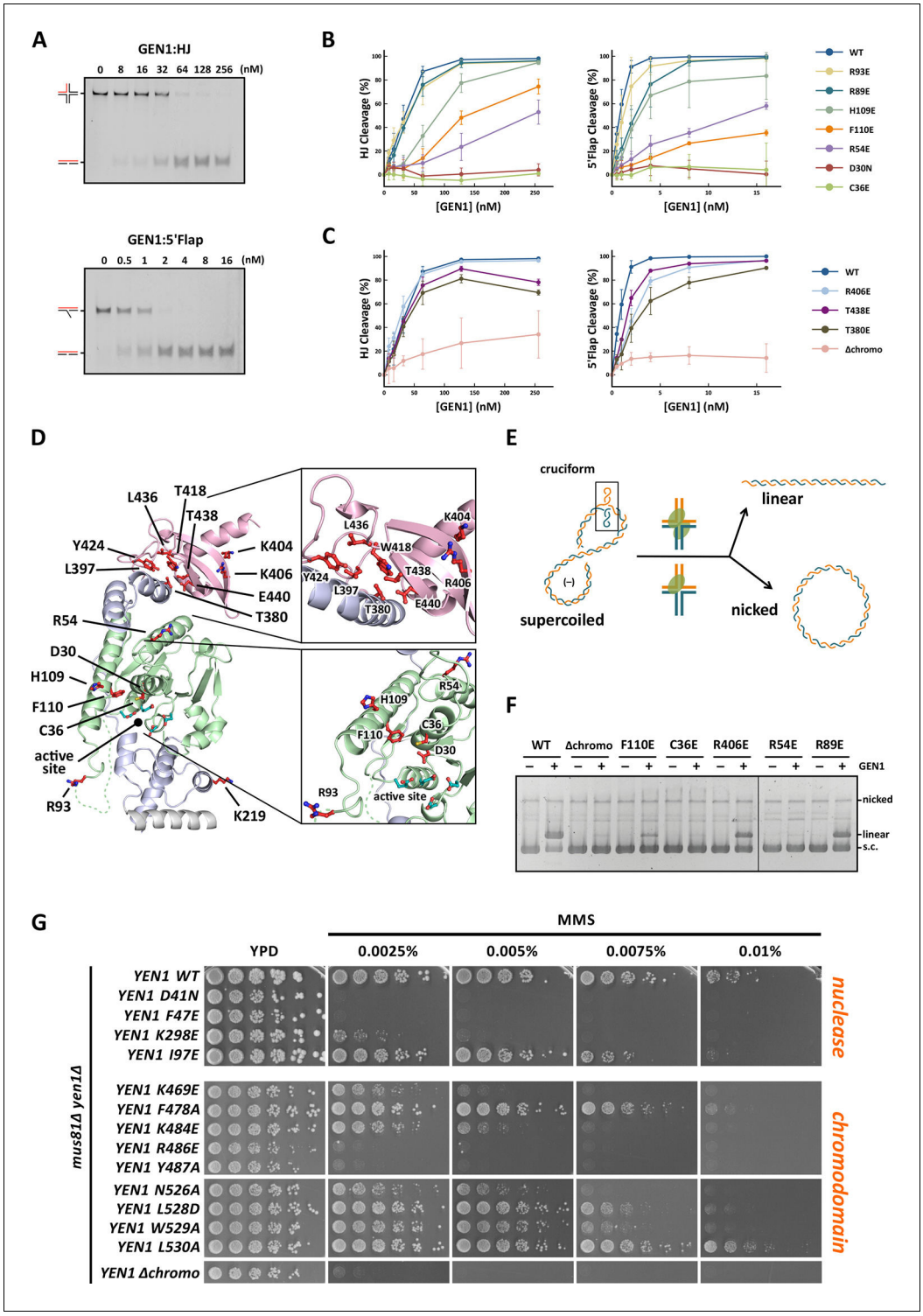


Figure 5. Functional analysis of GEN1. (**A**) Nuclease activity of GEN1 with HJ and 5′flap DNA. 40 nM 5′ 6FAM-labeled substrates were mixed with indicated amounts of GEN1. Reactions were carried out at 37°C for 15 min, products were separated by native PAGE and analyzed with a phosphoimager. ***Figure 5—source data 1*** gives the sequences of DNA oligos used in biochemical assays and ***Figure 5—source data 3*** shows activity measurements. (**B**) Quantification of nuclease assays of wild type GEN1 and variants with mutated residues located at the protein-DNA interfaces. Percentage of cleavage was plotted against the enzyme concentration. Error bars depict the standard deviation calculated from at least three independent experiments. ***Figure 5—figure supplement 1***

Figure 5 continued on next page

Figure 5 continued

shows representative gels from the PAGE analysis. (**C**) Quantification of nuclease assays of wild type GEN1 and variants with mutated residues located at the chromodomain. Error bars depict the standard deviation calculated from at least three independent experiments. ***Figure 5—figure supplement 2*** shows representative gels from the PAGE analysis. (**D**) GEN1 mutations used in this study. Locations of human GEN1 mutations used in biochemical assays and corresponding residues in yeast MMS survival assays are highlighted in red. Active site residues E134, E136, D155, D157 are marked in turquoise. (**E**) Schematic of the cruciform plasmid cleavage assay. A cruciform structure can be formed in plasmid pIRbke8^{mut}, which harbors an inverted-repeat sequence and is stabilized by negative supercoiling. Introducing two cuts across the junction point within the lifetime of the resolvase-junction complex yields linear products whereas sequential cleavage generates nicked products and the relaxed plasmid cannot be a substrate for the next cleavage. (**F**) Cruciform plasmid cleavage assay with different GEN1 variants. Plasmid pIRbke8^{mut} was treated with 256 nM GEN1 each and reactions were carried out at 37°C for 15 min. Supercoiled, linear and nicked plasmids were separated by native agarose gel electrophoresis and visualized with SYBR safe under UV light. (**G**) MMS survival assays with yeast y*en1* variants. The survival of *yen1* mutants was tested under a *yen1*Δ *mus81*Δ background with indicated amounts of MMS. The top part shows mutations at GEN1-DNA interfaces and the bottom part mutations at the chromodomain (compare ***Figure 5—figure supplement 3*** for all controls and expression tests). ***Figure 5—source data 2*** gives a list of all yeast strains.
DOI: 10.7554/eLife.12256.013
The following source data is available for figure 5:

Source data 1. Oligonucleotides used in biochemical assays.
DOI: 10.7554/eLife.12256.014
Source data 2. Yeast strains used for MMS survival assays.
DOI: 10.7554/eLife.12256.015
Source data 3. In vitro activity measurements of different GEN1$^{2\text{-}505}$ variants.
DOI: 10.7554/eLife.12256.016

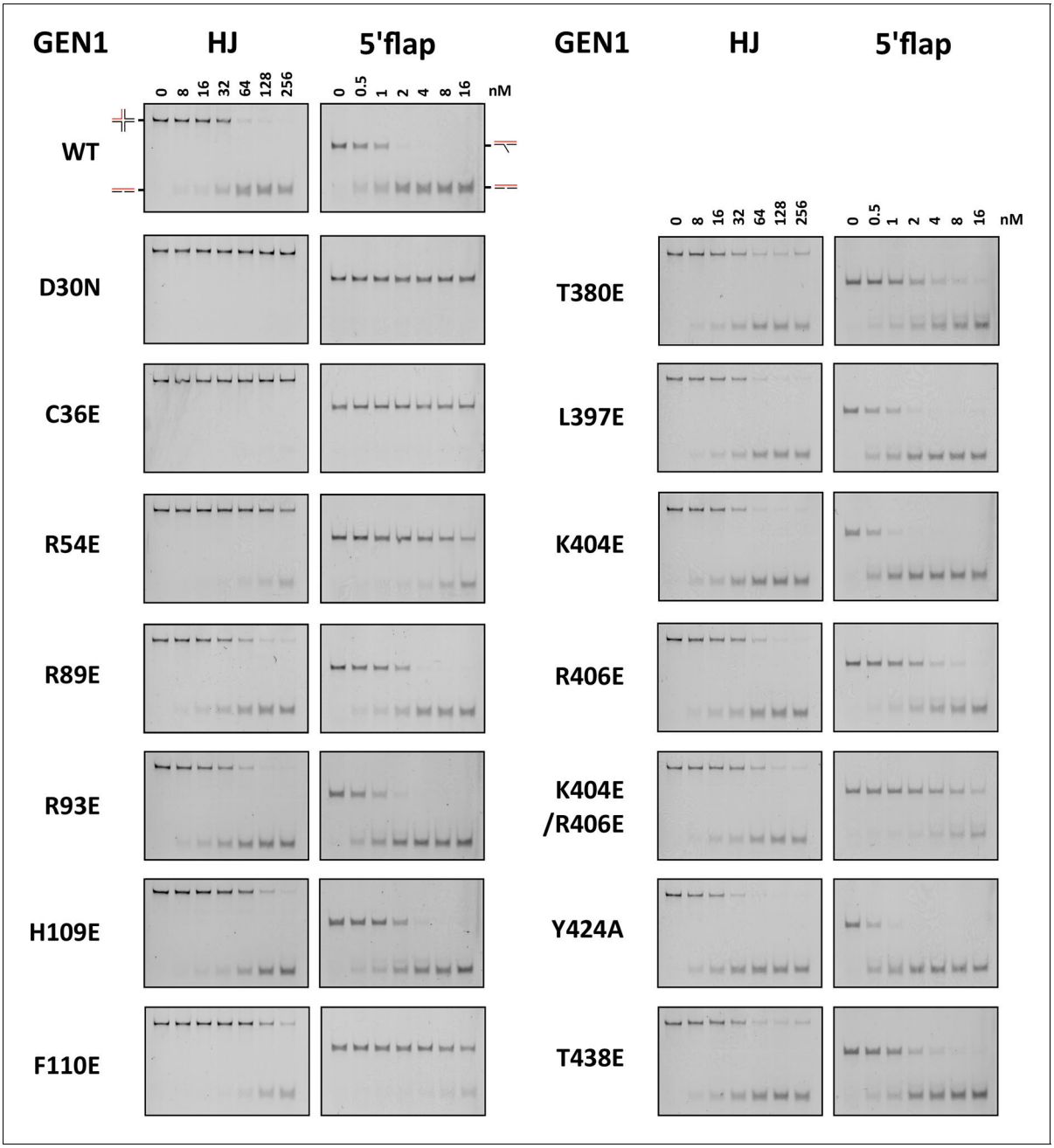


Figure 5—figure supplement 1. DNA cleavage assays of different GEN1 mutations. All GEN1$^{2-505}$ mutations were generated by site-directed mutagenesis and purified with the same procedure. Experiments were repeated three times and a representative gel picture is shown for each protein variant in ***Figure 5***.

DOI: 10.7554/eLife.12256.017

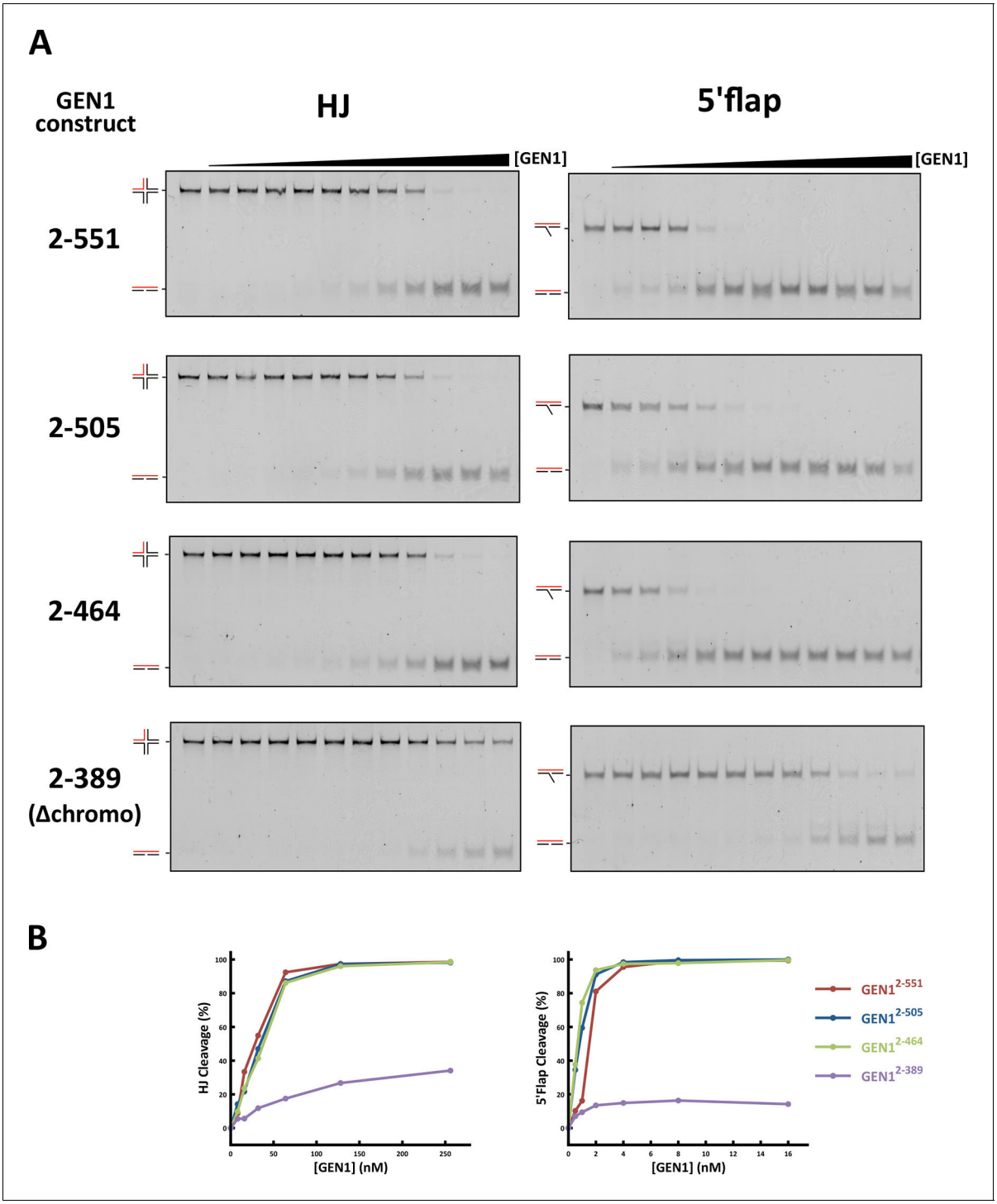


Figure 5—figure supplement 2. DNA cleavage assays of different GEN1 fragments. (**A**) 5′ 6FAM labeled four-way junction or 5′flap DNA (40 nM) were mixed with varying concentrations of GEN1 truncations (0.25, 0.5, 1, 2, 4, 8, 16, 32, 64, 128, 256 nM, respectively). (**B**) Quantification of activity assays.
DOI: 10.7554/eLife.12256.018

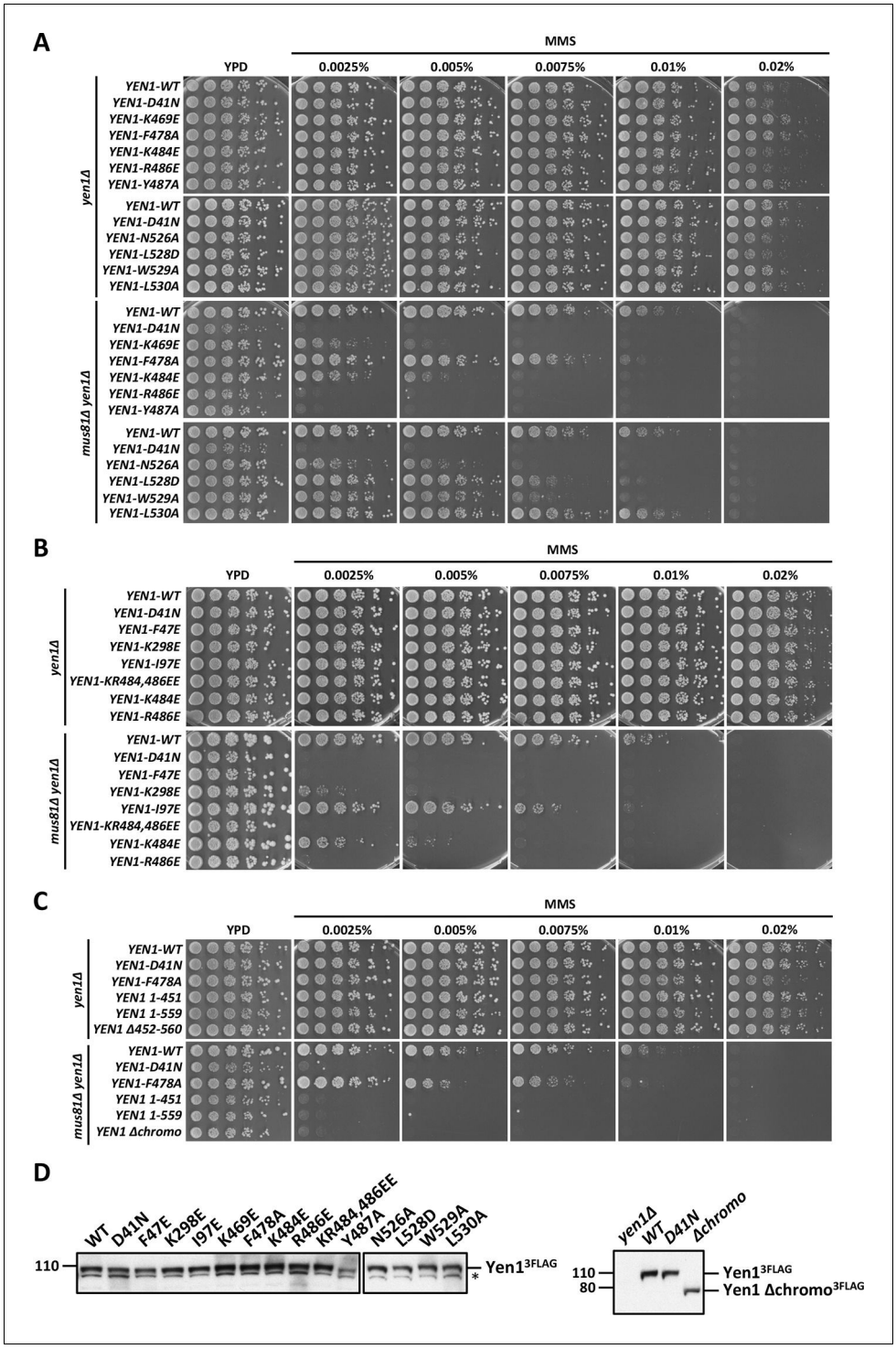


Figure 5—figure supplement 3. MMS survival assays with yeast *yen1* mutants. The survival of *yen1* mutants was tested in a *yen1*Δ or *yen1*Δ *mus81*Δ background with indicated amounts of MMS (compare ***Figure 5*** and ***Figure 5—source data 2***). Mus81 overlaps with Yen1 functionally, therefore *yen1*Δ knock-out strains are fully viable even in the presence of MMS, and hypersensitivity is only seen in the double knock-out. (**A**) Mutations in the chromodomain. (**B**) Mutations at protein-DNA interfaces. (**C**) *Yen1* truncations and chromodomain deletion. (**D**) Protein expression test (Western Blot analysis) of 3FLAG tagged Yen1 variants. Asterisk denotes a cross-reactive band.

DOI: 10.7554/eLife.12256.019

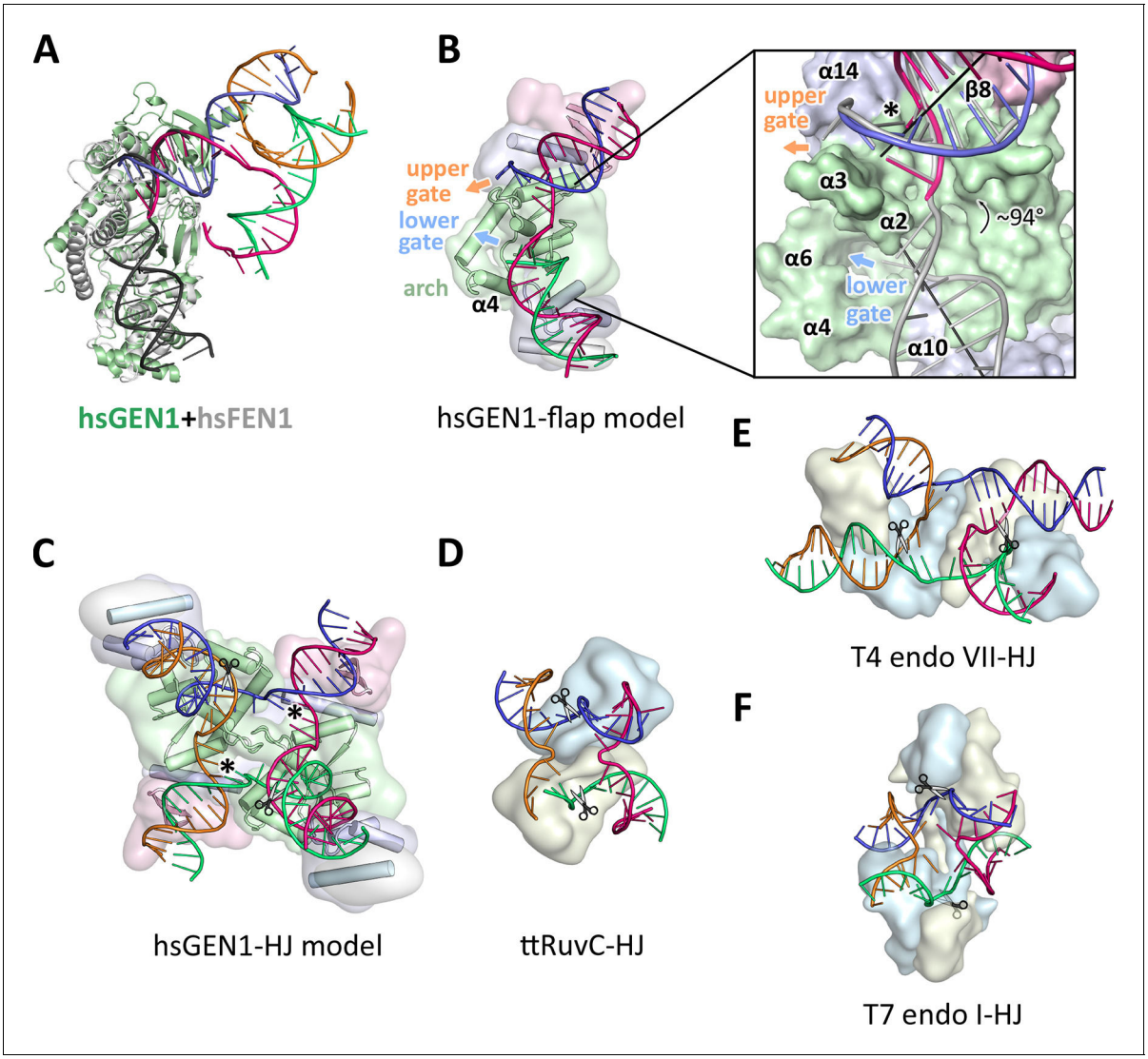


Figure 6. Substrate recognition features of GEN1. (**A**) Superposition of the protein part of the FEN1-DNA complex (PDB 3q8k, protein in gray, DNA in black) onto the GEN1-HJ complex (protein in green and the DNA strands in different colors). The FEN1-DNA aligns with the same register as the GEN1-DNA at the upstream interface. (**B**) Model for the recognition of a 5′ flap substrate by GEN1. The DNA was extended using the superimposition from **A**. Homology modeling suggests an additional helix α4 (disordered residues 79–92) forming an arch with helix α6. The protein is shown in a simplified surface representation with the same colors as in ***Figure 1*** and structural elements are highlighted. The insert shows a zoomed in view of the hydrophobic wedge with the modeled FEN1-DNA in gray. (**C**) Model for the dimerization of GEN1 upon binding to a HJ substrate based on the 5′ flap model in B. The monomers interlock via both arches (α4-α6) and the hydrophobic wedges (α2-α3) contact each other. (**D**) Structure of the *Thermus thermophilus* RuvC-HJ complex (PDB 4ld0). (**E**) Structure of the T4 endonuclease VII-HJ complex (PDB 2qnc). (**F**) Structure of the T7 endonuclease I-HJ complex (PDB 2pfj). Individual monomers are in surface representation, colored in light blue and beige, respectively. DNA strands are shown as ladders in different colors.

DOI: 10.7554/eLife.12256.020

Chapter 3 –
Two-mode DNA Recognition by Human Holliday Junction Resolvase GEN1

This chapter is a follow-up study based on the analysis of the human GEN1 crystal structure. This study provides biochemical evidence that the GEN1 helical arch is the key molecular module for substrate discrimination. Particularly, the helical arch is important for the highly efficient 5′ flap cleavage. In addition, the complexes of GEN1-HJ and GEN1-5′ flap were successfully reconstituted and analyzed, confirming that GEN1 is in monomeric form when bound to 5′ flaps but dimerizes upon binding to HJs. Taken together, this study supports the proposal that human GEN1 utilizes different mechanisms to recognize HJs and 5′ flaps, and implying versatile roles of GEN1 in the cells.

This study was conducted under the supervision of Dr. Christian Biertümpfel. Lee, S.-H. participated in the conception and design of the project, data acquisition, data analysis and interpretation, preparation of the first draft and revising the manuscript. Detailed author contributions are included in the attached article.

Two-mode DNA Recognition by Human Holliday Junction Resolvase GEN1

Shun-Hsiao Lee[1,2], Claire Basquin[2] and Christian Biertümpfel[1,2]*

[1]Molecular Mechanisms of DNA Repair

[2]Department of Structural Cell Biology

Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

*For correspondence: biertuempfel@biochem.mpg.de

Abstract

GEN1 is a specialized Rad2/XPG nuclease that resolves Holliday junctions (HJs) in eukaryotes. The resolution of HJs requires timely dimerization of GEN1 to introduce symmetrical incisions in order to maintain genome integrity. 5' flaps, replication forks and splayed arms are reported as GEN1 substrates as well. However, the mechanisms of resolving those structures are not well understood. A series of structure-based mutagenic studies were performed to delineate GEN1's substrate recognition. Results show that a preserved helical arch is the critical element to facilitate 5' flap cleavage, suggesting a similar biochemistry to other flap endonucleases. In addition, we identified a positively charged cluster located C-terminal of the chromodomain, which facilitates additional protein-DNA interactions. This interaction is in general required for proper cleavage under physiological salt concentrations. Finally, our static light-scattering and small-angle X-ray scattering analyses confirm that GEN1 functions as a monomer when recognizing 5' flaps and it dimerizes on Holliday junctions. We propose a model that GEN1 first scans DNA unspecifically via its positively charged cluster and then, its helical arch adopts a conformation that can discriminate various DNA structures in an induced-fit like manner. This mechanism ensures a fast scanning mode while only specific complexes are cleavage competent. Furthermore, it explains how the dimerization-triggered catalysis of Holliday junctions is bypassed when recognizing flap substrates.

Introduction

Homologous recombination is an important tool to repair DNA double-stranded breaks (DSBs) in mitotic cells as it ensures genome integrity in an error-free fashion (1,2). It is also essential for meiosis, which facilitates chromosome alignment and introduces gene diversity to the offspring (3,4). DSBs are first recognized and resected to produce coated 3' overhangs. These nucleoprotein complexes invade homologous DNA duplexes and form a D-loop intermediate, which provides a correct template to restore the genetic information. If a D-loop is stabilized and the newly synthesized DNA recaptures the original duplex, a double Holliday junction is created (5). These joint molecules have to

be eliminated in order to avoid improper DNA segregation. In human cells, two pathways have been identified that are in charge of joint molecule removal during homologous recombination. In the “dissolution” pathway, the BTR complex comprising BLM helicase, Topoisomerase IIIα, RMI1 and RMI2 “dissolves” double Holliday junctions via convergent branch migration and single-stranded passage activity (6,7). The reaction yields strictly non-crossover products hence it is the preferred pathway in somatic cells avoiding sister chromatid exchanges (SCEs) and loss of heterozygosity (8). The alternative “resolution” pathway is mediated by structure-selective endonucleases (SSEs) that generate both crossover and non-crossover products (9). HJ resolvases SLX1-SLX4 and MUS81-EME1 (SLX-MUS) work cooperatively on joint molecules and create mostly non-ligatable products (10-15). The activity is modulated by CDK1-mediated phosphorylation during early stages of mitosis. Independently, the 5’ endonuclease GEN1 resolves Holliday junctions in a symmetrical and concerted fashion leading to ligation-competent products (16-18). The regulation of GEN1’s catalytic activity varies among species. Human GEN1 carries a nuclear export signal and it is excluded from the nucleus, therefore has no access to substrates until nuclear break-down in prometaphase (19). However, its yeast homolog Yen1 is activated in anaphase by Cdc14-mediated dephosphorylation (20-22). Both nucleolytic systems are tightly controlled throughout the cell cycle (23,24). The current view is that SLX-MUS is the dominant nuclease for HJ resolution, and GEN1 plays a safe-guarding role to ensure that all joint molecules are fully resolved before chromosome segregation (25).

GEN1 belongs to the Rad2/XPG family of structure-specific nucleases that have a conserved nuclease core and participate in various DNA repair pathways (16,26). Interestingly, GEN1 is the sole member that is able to form dimers upon HJ binding and cleaves its substrate cooperatively. It has been shown that GEN1 can dimerize on HJ and make symmetrical incisions one nucleotide away from the junction point in a nick and counter-nick mechanism (16-18). The cleavage has weak sequence specificity in between two G residues near T-rich region (27). Dimerization-coupled cleavage is an elegant way to ensure that four-way junctions are correctly resolved into two nicked duplexes, which in turn are resealed by endogenous ligases. Since this mechanism is analogous to the prototypical *E. coli* RuvC resolvase, GEN1 is considered as a classical Holliday junction resolvase in eukaryotes. Even though the crystal structures of human

GEN1 catalytic core as well as its homolog from *Chaetomium thermophilum* have been solved, the mechanism of protein dimerization is still elusive (28,29).

Biochemically, GEN1 has a broad spectrum of substrate specificities, besides HJ resolution, cleavage of 5' flaps, replication forks and splayed arms have been reported (17,30,31). A recent study has shown that the yeast homolog, Yen1, removes replication intermediates in Dna2-deficient cells, suggesting that Yen1 has also the ability to recognize different substrates *in vivo* (32). Despite extensive studies on GEN1's HJ resolution activity, it is not clear how GEN1 discriminates between HJs, 5' flaps and replication forks. Intuitively, GEN1 is able to cleave less restraint structures like 5' flaps and replication forks in monomeric form as the reaction requires only one active site positioned at the scissile phosphate. In contrast, studies on HJ resolution indicate that GEN1's catalytic activity is inhibited until a GEN1 dimer is fully assembled around a HJ. Therefore, it is an open question that whether GEN1 uses the same recognition mechanism for 5' flap, replication fork and HJ substrates and what triggers its nuclease activity.

5' flap cleavage by Rad2/XPG nucleases is well characterized by studies on the GEN1 paralog FEN1. The α4, α5 and α6 helices of FEN1 undergo a "disorder-to-order" transition upon substrate binding and form a helical arch with a cap threading the 5' single-stranded flap through the tunnel and thereby, positioning the substrate for cleavage (33,34). This helical arch arrangement is also conserved in another Rad2/XPG nuclease, EXO1 (35). In contrast, the arch region in GEN1 comprises only helices α4 and α6 without a cap and an extended α6 was observed in the crystal structure of human GEN1 (28). Our previous study has suggested that the extended α6 helix might serve as a molecular lever to discriminate substrates that allows HJ accommodation and provides additional interactions to stabilize flexible substrates like 5' flap (28).

Here, we report a series of mutagenesis studies focusing on the arch region. The results indicate that the disordered loop on the arch is the critical element for substrate recognition. Since the arch mutations present a different degree of preference towards HJ or 5' flap substrates, we propose that the helical arch is an activity booster for 5' flap cleavage. Our hydrodynamic analysis shows, consistent with previous biochemical studies, that GEN1 dimerizes upon binding to HJ. Interestingly, GEN1 remains

monomeric when associating with 5' flaps, demonstrating that GEN1 uses different mechanisms to disentangle various DNA structures, and the dimerization-triggered cleavage is bypassed while resolving flap structures.

Materials and methods

Chemicals and reagents

Chemicals and reagents were purchased from AppliChem (Darmstadt, Germany), Carl Roth (Karlsruhe, Germany), Sigma Aldrich (Taufkirchen, Germany) unless indicated otherwise. All oligonucleotides were obtained from Eurofins Genomics (Ebersberg, Germany).

Cloning of expression plasmids

All expression constructs and truncations were cloned as previously reported by ligation-independent cloning (LIC) or site-directed mutagenesis (28). The Δloop variant $GEN1^{1\text{-}464\Delta L}$ was designed with two cleavage sites for HRV 3c protease (LEVLFQ|GP) before position 245 and after position 307 removing 245-307 and leaving additional residues from the cleavage.

Protein purification

Human $GEN1^{1\text{-}505}$ and mutants thereof were expressed and purified as previously described with slight modifications (28). Briefly, proteins were expressed in *E. coli* BL21(DE3) pRIL cells (MerckMillipore, Darmstadt, Germany) in 4 l TB medium and induced with 0.2 mM IPTG at 16°C overnight. Cell were harvested by centrifugation at 8 950 g and resuspended in lysis buffer (1x phosphate-buffered saline supplemented with 500 mM NaCl, 1 mM EDTA, 2 mM DTT, 0.3 µM aprotinin, 1 µM leupeptin, 1 µM pepstatin A and 2 µM AEBSF, 10% (v/v) glycerol). Cells were disrupted by sonication and debris removed by centrifugation at 75 600 g for 45 minutes. The cleared lysate was loaded onto a 5 ml cOmplete His-Tag Purification column (Roche Diagnostics, Mannheim, Germany) and washed with 20 column volumes of Ni-wash buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 10% (v/v) glycerol and 7.5 mM imidazole). Proteins were eluted with

Ni-B buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 10% (v/v) glycerol and 300 mM imidazole), C-terminal tag cleaved overnight, diluted with SP-A buffer (20 mM Tris-HCl pH 7.5, 10% (v/v) glycerol and 1 mM EDTA) to a final concentration of 150 mM NaCl and loaded onto a 5 mL HiTrap SP HP cation exchange column (GE Healthcare, Freiburg, Germany). Proteins were eluted with linear KCl gradient from 150-450 mM. Peak fractions were concentrated and injected to a HiLoad 16/600 Superdex 200 column (GE Healthcare) pre-equilibrated with SEC buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 5 % (v/v) glycerol, 0.1 mM EDTA, 2 mM TCEP). Proteins were concentrated to 10-35 mg/ml, flash-frozen in liquid nitrogen and stored at -80°C until further use. For the purification of $GEN1^{1\text{-}464\Delta L}$, the same procedures were applied, except that an extra 10 µg/ml 3C protease was included in the overnight cleavage to remove the flexible loop.

Annealing of DNA substrates

In order to prepare DNA substrates, complementary oligonucleotides were mixed in equimolar ratios in annealing buffer (10 mM HEPES-KOH pH 8.5, 50 mM KCl, 0.1 mM EDTA), heated to 85˚C for 5 min, slow-cooled to room temperature over 90 min and stored at -20˚C until further use. The oligonucleotides used in this study were listed in Supplementary Table 1.

Nuclease Assays

Cleavage reactions were set up in a final volume of 10 µl containing indicated amounts of enzyme with 40 nM 6FAM-labeled substrates in 20 mM Tris-HCl pH 8.0, 50 ng/µl bovine serum albumin (BSA), 1 mM DTT. Extra KCl was included in salt titration assays as indicated. Reactions were initiated by adding $MgCl_2$ at a final concentration of 5 mM, incubated at 37°C for 15 minutes and terminated by adding 15 mM EDTA and 0.3% SDS. Reactions were further deproteinized by adding 1 mg/ml proteinase K at 37°C for 15 minutes. Samples were mixed with 1.2 µl of 10x loading buffer, 10 µl loaded onto an 8% 1x TBE native polyacrylamide gel (19:1) and run at 100 V for 50 minutes. The fluorescence signal was detected with a Typhoon 7000 phosphoimager (GE Healthcare) quantified with ImageJ (36) and visualized with Prism (GraphPad, La Jolla, CA, USA).

Substrate binding assays

Samples were set up in 10 µl reactions containing indicated amounts of enzyme with 40 nM 6FAM-labeled DNA substrates in 20 mM Tris-HCl pH 8.0, 50 ng/µl bovine serum

albumin (BSA), 1 mM DTT and 5% (v/v) glycerol. Reactions were incubated on ice for 10 minutes and immediately loaded onto a 6% 0.5x TBE polyacrylamide gel (29:1). An additional 1% (v/v) glycerol was included in both gel and running buffers, which improved the resolution of bands. Electrophoresis was carried out on ice at constant 2W for 30 minutes (double-stranded DNA) or an hour (HJ and 5' flap). The fluorescence signal was detected with a Typhoon 7000 phosphoimager and analyzed with ImageJ and Prism (see also Nuclease Assays).

Limited Proteolysis

Human GEN1$^{1\text{-}505\ D30N}$ was used for limited proteolysis. 10 µl reactions containing 0.6 mg/ml protein were mixed with equimolar DNA in 70 mM Tris-HCl pH 8.0, 50 mM NaCl and 10% (v/v) glycerol. Reactions were initiated by adding 3 µl of 0.1 mg/ml GluC protease on ice, and terminated by 5 µl of 3x SDS loading buffer in indicated time points. To determine the protease cleavage positions, same reaction without DNA was scaled up and injected into a Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated with 20 mM Tris-HCl pH 7.5, 100 mM NaCl and 5% (v/v) glycerol. Peak fractions were collected and the molecular weights of proteolytic products were determined by ESI-TOF with micrOTOF (Bruker, Bremen, Germany).

Cruciform-cutting Assay

The cruciform plasmid pIRbke8^{mut} is a gift from Stephen West's lab. 50 ng/µl plasmids were mixed with 20 mM Tris-HCl pH 8.0, 50 mM potassium glutamate, 5 mM $MgCl_2$, 50 ng/µl BSA and 1 mM DTT and pre-warmed at 37°C for 1 hour to induce the cruciform structure (37). Reactions were initiated by adding indicated amount of GEN1 and incubated at 37°C for 15 minutes and terminated as DNA cleavage assays. Products were separated by 1% 1x TBE native agarose gel and stained by SYBR safe after electrophoresis.

Static light-scattering (SLS) analysis

GEN1$^{1\text{-}464\Delta L}$ alone, a GEN1$^{1\text{-}464\Delta L}$-5' flap mixture at a molar ratio of 1:1 or a GEN1$^{1\text{-}464\Delta L}$-HJ-mixture at a molar ratio of 2:1 were run on a Superdex 200 5/150 GL column on an ÄKTAmicro system (GE Healthcare) coupled to a Viscotek TDA302 detector (Malvern, Herrenberg, Germany) in buffer containing 20 mM Tris-HCl pH 7.5, 60 mM KCl, 2 mM $CaCl_2$, 2% glycerol and 2 mM TCEP. 20 µl samples were injected at a concentration of 3

mg/ml (GEN1$^{1\text{-}464\Delta L}$-5' flap and GEN1$^{1\text{-}464\Delta L}$-HJ) or 10.5 mg/ml (GEN1$^{1\text{-}464\Delta L}$ alone). Bovine serum albumin was used as a standard and the refractive index increment (dn/dc) was set to 0.180 ml/g for calculations. Data was analyzed using the OmniSEC 4.5 software (Malvern).

SEC-SAXS

All data was collected at the European Synchrotron Radiation Facility (ESRF) beamline BM29 with PILATUS 1M detector. Samples were prepared at room temperature with the following conditions: 8 mg/ml GEN1$^{1\text{-}464\Delta L}$, GEN1$^{1\text{-}464\Delta L}$ at 4 mg/ml with 5' flap substrate at a 1:1 molar ratio, and GEN1$^{1\text{-}464\Delta L}$ at 4 mg/ml with Holliday junctions at a 2:1 molar ratio. Each sample was injected into a Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated with 20 mM Tris-HCl pH 7.5, 60 mM KCl, 2 mM $CaCl_2$, 2% glycerol and 2 mM TCEP at room temperature. Frames were collected at 1 Hz at flow rate 0.5 ml/min. The data was processed using the ATSAS package with the standard procedures (38). Data reduction was performed with the PRIMUS program (39). The Guinier approximation of I_0 and R_g values were calculated by choosing the low angle points at a linear range (q_{max}*R_g < 1.3). The real space R_g, D_{max}, and the Porod volumes were derived from the pairwise distance distribution functions, P(r), by the program GNOM (40). Molecular weights were calculated from the raw data with the program SAXS MoW2 (http://saxs.ifsc.usp.br/) (41). To reconstruct *ab initio* shapes, the scattering data from each sample was processed by DAMMIF assuming P1 symmetry for GEN1$^{1\text{-}464\Delta L}$ and the GEN1$^{1\text{-}464\Delta L}$-5' flap complex, P2 symmetry for the GEN1$^{1\text{-}464\Delta L}$-HJ complex. Ten initial dummy-residue models were further averaged by DAMAVER and refined by DAMMIN (42-44). Atomic models were generated from PDB code 5t9j (28) and manually placed into the filtered envelopes from DAMMIN followed by rigid-body fitting in Chimera (45). The fitness of each atomic model against scattering data was evaluated by the program CRYSOL (46).

Results

Structure comparison of GEN1 homologs

The crystal structure of human GEN1 (hsGEN1) reveals a highly conserved Rad2/XPG nuclease core (Figure 1 A, green and blue domains), suggesting a similar catalytic mechanism as other family members (28,47). However, the conformations of the arch region (residues 78-125 in hsGEN1) are rather divergent (Figure 1B) and a sequence alignment shows a low conservation within this region (Figure 1C). Interestingly, the arch region has been shown to undergo a disorder-to-order transition in FEN1 and EXO1 upon substrate binding. The helices α4 and α5 together with α6 create a helical arch that is instrumental for substrate positioning (33-35). This two-helix clamp is missing in GEN1, and the arch region is substituted by the extended helix α6 and connecting to a shorter disordered loop. This specialized conformation is strikingly similar to the one of bacteriophage T5 flap endonuclease (T5FEN, Figure 1B and S1) (48-50). A recent structural study of T5FEN identified two structural states of its arch region: an open state, in which a stretch of residues lost the secondary structure creating a large loop, and a closed state, in which residues 84-92 form an α-helix thereby narrowing the tunnel and only allowing single-stranded DNA to pass through (48). The close sequence similarity of the loop region in human GEN1 (compare Figure 1C) suggests the same conformational adaptability as in T5FEN. Interestingly, this arch region is missing in the structure of *Chaetomium thermophium* GEN1 (ctGEN1). It has a shorter helix α6 and the disordered loop does not seem to be long enough to create a tunnel for single-stranded DNA (Figure 1B) (29). Corresponding to the missing arch, ctGEN1 was also reported as a strict Holliday junction resolvase that only recognizes four-way junction and is not able to cut 5' flap and replication fork DNA, supporting the hypothesis that the human GEN1 α6 plays a critical role for substrate recognition (51). Taking these observations and biochemical analyses together, the dimerization feature of GEN1 and preserved structural elements enable human GEN1 to recognize versatile DNA structures and points to a critical role of the arch region.

The GEN1 arch is required to facilitate 5' flap cleavage

The GEN1 arch is highly positively charged and several basic residues face towards the arch tunnel (Figure 2A). They are positioned to interact with the negatively charged

DNA backbone and contribute to substrate recognition. To examine the function of these residues, we generated four alanine mutations (K99A, K104A, R107A and K111A) to deplete positive charges (Figure S2). We used C-terminally truncated GEN1 ($GEN1^{1-505}$) as a model to test substrate cleavage as it is fully functional and has been proved to have the same catalytic efficiency comparing to the full-length protein (18,28). The results show that the mutations of K99A, K104A and K111A have the same activity as the wild-type enzyme, but R107A has a lower activity on both, Holliday junction and 5' flap cleavage (Figure 2B). Interestingly, arginine 107 is located on the top part of the arch tunnel and the side chain points towards the center. The effect is more pronounced with 5' flap substrates compared to Holliday junctions, supporting the hypothesis that the single-stranded flap is going through the arch tunnel.

To identify the key part of the arch for substrate recognition, we created four deletions (Δarch 1 - Δarch 4) centering around residue F110 which has been previously found to be critical for substrate binding (Figure 2A) (28). In the case of FEN1 the corresponding residue V133 is critical in stabilizing 5' flap DNA by orienting the -1 nucleotide for catalysis (33). All deletions were successfully purified and have similar stabilities as wild-type protein except Δarch 1 (deletion 80-115), which is prone to precipitate at low salt conditions and is catalytically dead. The other deletions are able to cleave DNA and have a different degree of substrate preference (Figure S2). The Δarch 2, (deletion 80-107) removing the disordered loop and the first two helical-turns of α6, shows a significant reduction in activity. The activity is about 63-fold reduced for 5' flap and about 6-fold for HJ DNA (judged from 50% catalytic rate) (Figure 2C). Holliday junction cleavage is only slightly weaker than the R107A mutant, suggesting that the activity loss of Δarch 2 is mainly contributed by disrupting R107. In contrast, the Δarch 2 has a much weaker activity on 5' flaps compared to R107A, about ten times more enzyme is needed to reach to the same level of 5' flap cleavage. This bias indicates that the helical arch plays a critical role for substrates discrimination. To further decipher the role of the helical arch, we tested the Δarch 3 (deletion 80-94) deleting only the disordered loop part. Surprisingly, the activity of Δarch 3 has nearly the same efficiency as Δarch 2 on both Holliday junction and 5' flap substrates, demonstrating that the disordered loop is the key element for 5' flap recognition. Consistent with this observation, the Δarch 4 (deletion 92-105), which only deletes the first two turns of α6 and part of the linker, has a comparable activity on Holliday junction as wild-type protein and shows a mild effect

on 5’ flap DNA. Taken together, the deletion of the helical arch negatively impacts the activity of GEN1 and it has a bias towards 5’ flap over Holliday junction substrates, implying that GEN1 uses different mechanisms to recognize various DNA structures. More precisely, we identified the disordered loop on the helical arch as a critical element for 5’ flap recognition. As GEN1 constructs with the loop deletion show only slightly reduced the cleavage activity for HJs and significant impairment for 5’ flaps, we hypothesize that the disordered loop is an activity booster for flexible substrates. It is noteworthy that the helical arch is not involved in protein dimerization as all functional arch deletions create only linear products in the cruciform plasmid assays (Figure S3). The mechanism is likely similar to T5FEN that the 5’ single–stranded DNA pass through the arch and the disordered loop becomes ordered to assist substrate stabilization.

GEN1’s positively charged cluster participates in DNA interactions

It has been shown that GEN1’s chromodomain is essential for proper substrate recognition and cleavage (28). Sequence analysis shows a cluster of positively charged residues located C-terminal of the chromodomain. This part was essential for crystallization but disordered in the crystal structure (28) and the functional relevance of this region has not been tested yet. A sequence alignment shows that this basic region is not conserved among different species but it has a same tendency of harboring positively charged residues (Figure 3A). To address the function of this region, we first purified a truncated form of GEN1 comprising the catalytic core from residue 1-464 without the positive charge cluster. Electrophoretic mobility shift assays (EMSAs) were carried out and the results showed that $GEN1^{1\text{-}464}$ has a weaker DNA binding activity compared to $GEN1^{1\text{-}505}$ (Figure 3B). This lower affinity is seen for Holliday junction, 5’ flap as well as double-stranded DNA indicating that this DNA interaction mode is non-specific. This additional interaction is further supported by limited proteolysis. Comparing the proteolysis pattern of GEN1 protein alone and a GEN1-DNA mixture, we identified region 465-505 as resistant to GluC protease cleavage when DNA is present, indicating that this region interacts with DNA thus, is not accessible for protease digestion (Figure 3C-D and S4).

To further identify the role of the positively charged cluster, we tested substrate cleavage of $GEN1^{1\text{-}505}$ and $GEN1^{1\text{-}464}$ side-by-side under different salt concentrations (Figure 3E and 3F). Both constructs have the same cleavage activity on Holliday junction

and 5' flap DNA at low salt concentrations confirming that the positively charged cluster does not directly participate in catalysis or substrate discrimination. However, the activity of GEN1$^{1-464}$ is more sensitive to salt conditions. At 150 mM KCl and 60 nM protein, GEN1$^{1-505}$ is able to cleave around 90% of Holliday junction substrate, but the same amount of GEN1$^{1-464}$ can only cleave about 50%. The salt sensitivity also seen in 5' flap cleavage. At 100 mM KCl, approximately 65% of 5' flap is cleaved by 1 nM of GEN1$^{1-505}$, but GEN1$^{1-464}$ only shows 10% cleavage. These results demonstrate that the positively charged cluster has a function to promote the enzymatic activity despite its non-specific DNA interaction. Presumably, it serves as a molecular tether that facilitates GEN1 searching for DNA substrates.

GEN1 forms a stable dimer on Holliday junctions and remains monomeric on 5' flaps

It is difficult to reconstitute the functional GEN1-HJ complex for biochemical studies, as human GEN1 tends to be unstable when interacting with DNA. To optimize protein stability, we tested different GEN1 truncations and deletions in order to search for the minimum core that forms stable complexes with DNA substrates and still retains enzymatic activity. Region 245-307 was disordered in the crystal structure of human GEN1$^{1-505}$ (28). To assess the role of this region, we introduced two HRV 3C protease cleavage sites flanking this loop and removed it during purification. The recombinant protein is named GEN1$^{1-464\Delta L}$. Biochemical analysis showed that GEN1$^{1-464\Delta L}$ is still fully functional and retains the same DNA binding activity as wild-type protein, indicating that this disordered loop is dispensable for catalysis (Figure S5). Surprisingly, the GEN1$^{1-464\Delta L}$ is highly stable with DNA substrates allowing a large-scale biochemical analysis. The removed loop contains nine cysteine residues, which likely contribute to the aggregation behavior of GEN1. Our static light-scattering analysis showed that GEN1$^{1-464\Delta L}$ is predominantly a 48 kDa species in solution, agreeing with previous studies that GEN1 is a monomer in solution (17,18). Moreover, we detected a 69 kDa species from the GEN1-5' flap complex and a 136 kDa species from the GEN1-HJ mixture, confirming that GEN1 cleaves 5' flap DNA as a monomer and it forms a dimer upon Holliday junction binding (Figure 4A).

Solution structure of GEN1 in complex with DNA substrates

To observe the higher order conformational arrangement of GEN1 upon substrate-binding, the stable GEN1$^{1\text{-}464\Delta L}$ construct was further characterized by size-exclusion chromatography coupled with in-line small-angle X-ray scattering (SEC-SAXS). This approach is very powerful to characterize proteins and protein complexes in solution as sample homogeneity is increased by gel filtration and aggregates as well as unbound species are separated before applying to X-ray scattering (52). The radius of gyration (R_g) of GEN1$^{1\text{-}464\Delta L}$ is 27.4 Å with a maximum distance (D_{max}) of 93.4 Å, and the pairwise distance distribution function, P(r), suggests an extended structure which is in good agreement with the dimensions of the GEN1$^{1\text{-}505}$ crystal structure. The atomic model fits well into an *ab initio* solution structure with a χ^2 value of 1.03 (Figure 4B-E). A slightly larger envelope was calculated from GEN1$^{1\text{-}464\Delta L}$ in complex with a 5' flap structure. The R_g and the D_{max} were determined to 31.3 Å and 98.8 Å, respectively. The molecular weight calculated from the SAXS data is 84.2 kDa, which is slightly larger than the calculated mass of 66.8 kDa for the complex (Figure 4C). Further, the expanded volume of the envelope enabled us to place a model of GEN1$^{1\text{-}464\Delta L}$ in complex with 5' flap DNA with a χ^2 value of 1.79 (Figure 4D and E). In contrast, the GEN1$^{1\text{-}464\Delta L}$-HJ complex showed a significantly larger structure than the 5' flap complex with a R_g of 44.3 Å and a D_{max} of 153.6 Å. The calculated molecular weight of 152.1 kDa fits well to that theoretical mass of 131.2 kDa. The differences between calculated and measured masses for the GEN1-DNA complexes likely results from an extended hydration shell around the DNA, which contributes additional scattering. The square-shaped *ab initio* GEN1$^{1\text{-}464\Delta L}$-HJ solution structure fits well with a model for the complex with a χ^2 value of 2.98 (Figure 4B-E). It is noteworthy that two of the HJ arms from our model are partially out of the *ab initio* structure. This deviation can be due to a high flexibility of the DNA arms or because of an inaccuracy of the angles of the DNA arms of our hypothetical model. Nevertheless, the parameters from solution studies strongly support that GEN1 is able to dimerize on HJs. In summary, the SAXS analysis of GEN1-DNA complexes provides direct evidence that human GEN1 remains in a monomeric state with 5' flap substrates while dimerizing upon binding to HJs. The precise structure of the GEN1-HJ complex has to be further studied by high-resolution techniques.

Discussion

The GEN1 protein is considered as the “classical” Holliday junction resolvase in eukaryotes, based on its biochemical features of forming a homodimer on four-way junctions and generating symmetrical incisions to yield ligatable products (16,17). However, the monomeric nature in solution makes it remarkably different from prokaryotic resolvases, which are obligate dimers in solution independent of binding to Holliday junctions. GEN1 follows a two-step mechanism for binding HJs, which is likely the explanation for its slow overall reaction rate on four-way junction substrates. The observation that GEN1 exclusively produces symmetrical nicks into HJs suggests that the rate-limiting step is the first binding event and that a single monomer is not cleavage-competent (18).

The slow cleavage rate disadvantage might not be harmful to the cell as the fact that GEN1 is not the only solution to joint molecules, and the relative inefficiency can be compensated by other resolvases (e.g. SLX1-SLX4, MUS81-EME1) and alternative pathways (e.g. dHJ dissolution). On the other hand, its monomeric nature could be beneficial to GEN1 to deal with various DNA structures. Besides Holliday junctions, 5’ flaps, replication forks and splayed arms have been reported as substrates of GEN1 (16-18). But how GEN1 differentiates these DNA structures is still illusive. Biochemical studies on Holliday junction resolution implies that the GEN1 activity is triggered by dimerization. Controversially, dimerization seems inefficient for other substrates as the cleavage of 5’ flaps, replication forks and splayed arms require only one catalytic center. Here we use 5’ flap as model substrate to compare the catalysis on Holliday junction, and identified the helical arch is a critical structural module for substrate discrimination.

Helical arch is the essential element for flap endonucleases, studies on human FEN1 protein have shown that its disorder-to-order transition is important for substrates recognition and positioning. Disrupting direct interactions between the arch and single-stranded flaps seriously affects enzymatic cleavage (53-55). The recent structure of the T5FEN-DNA complex provides direct evidence that the 5’ single-stranded DNA is able to thread through the tunnel created by the helical arch, and the residues on the arch offer interactions to guide the substrate to the correct position for subsequent cleavage (48).

Our structural analysis shows that human GEN1 has a highly similar arch region compared to the flap endonuclease in bacteriophage T5 (T5FEN). Structural similarities suggest that GEN1 also adopts this structural transition mechanism. It is noteworthy that the arch region is reduced in *Chaetomium* GEN1 and biochemical studies demonstrated that ctGEN1 is unable to cleave 5' flaps (51). This correlation further supports the hypothesis that the helical arch is important for GEN1 to recognize 5' flap substrates. Mutagenesis screening on the helical arch shows that Arg107 plays an important role in 5' flap cleavage by human GEN1. Even though mutating Arg107 to alanine affects the cleavage on both Holliday junction and 5' flaps, the effect is biased more towards flap structures. Around 5.5 times of enzyme is required to reach the same extent of 5' flap cleavage, comparing to around 3 times for Holliday junction. Since the Arg107 points directly toward the tunnel of helical arch, we suggest that this residue provides interactions to thread in and position single-stranded DNA. The residue likely cooperates with Phe110, which was previously found to be important for DNA recognition. The importance of the helical arch is also highlighted by the presented deletion studies. Deletion of the residues 80 to 94, which are invisible in the crystal structure, significantly reduces the activity on 5' flaps, indicating that the arch has a function to facilitate flap recognition. The arch deletions did not completely impair the activity. Instead, they seem to lose the substrate preference of 5' flap. Therefore, we propose that the helical arch provides specificity for 5' flap recognition and it and acts as an "activity booster".

Besides the flexible part of the helical arch playing an important role in substrate discrimination, we also identified another disordered region that has a function in DNA binding. Our results reveal that the positively charged cluster C-terminal of chromodomain interacts with DNA in a non-specific manner. Deleting the cluster did not affect the cleavage efficiency at low salt concentrations, indicating that the positive charge cluster does not directly involved in catalysis. However, titration experiments showed that this region is required for the efficient cleavage of both, Holliday junction and 5' flap DNA, near physiological salt concentrations. Presumably, GEN1 uses the positive cluster as a targeting module for searching DNA (Figure 5) and subsequent structure-specific interactions with correct substrates permit nuclease incision or in case of HJ substrates, a holding position for the binding of a second GEN1 molecule. Interestingly, human FEN1 also harbors a region rich in positive charges at the extended

C-terminus, which has been reported as a multiple protein interacting platform (56). We cannot rule out the possibility that there are additional functions of GEN1's positively charged cluster.

Biochemical and structural analysis of human GEN1-DNA complexes are hampered by the fact that they tend to precipitate at high concentrations. In order to capture protein-DNA complexes, we engineered a minimal GEN1 core ($GEN1^{1\text{-}464\Delta L}$) by removing a 63-residue extension of the 5'->3' exonuclease C-terminal domain (EXO). The shortened construct retains full catalytic activity and the solubility of protein-DNA complexes is drastically improved. It allowed us to perform a full range of static light-scattering and small-angle X-ray scattering experiments. The analysis showed directly that $GEN1^{1\text{-}464\Delta L}$ dimerizes on HJ and remains monomeric on 5' flap DNA. These results further confirm that GEN1 has different mechanisms to recognize different DNA structures, and dimerization is not required for 5' flap cleavage.

Previous crystal structure reveals that human GEN1 harbors several flexible regions around the nuclease core. In this study, we further examined these regions and discovered their function in substrates discrimination and DNA targeting. Our study highlights the versatile substrate recognition features of GEN1. The enzyme adopted a dimerization mechanism during evolution that ensures the HJ cleavage in a symmetric manner but also preserves the key elements for cleaving 5' flap substrates in monomeric form and bypass the dimerization-triggered catalysis. This versatility allows GEN1 to act on different substrates and it underlines the proposed role of GEN1 to be a last resort for cells for removing deadlocked DNA structures.

Acknowledgements

We thank Naoko Mizuno from the Max Planck Institute of Biochemistry for critical reading of the manuscript. The Biochemistry Core Facility of the Max Planck Institute of Biochemistry (MPIB) performed mass spectrometry analysis. We are grateful for the preliminary SEC-SAXS analysis of the original aggregation-prone GEN1 complexes performed by Clement Blanchet and Dimitri Svergun at the BioSAXS beamline P12 at

DESY/EMBL Hamburg. We also acknowledge the support by the staff of the SAXS beamline BM29 at ESRF Grenoble. Ralf Stehle and André Mourão helped with analysis of the SAXS data.

Author contributions

S-HL and CBier conceived the study and analyzed data. S-HL cloned, purified and analyzed proteins. CBasq performed SEC-SLS experiments. S-HL prepared a first draft of the manuscript and CBier revised it.

References

1. Heyer, W.D. (2015) Regulation of recombination and genomic maintenance. *Cold Spring Harb Perspect Biol*, **7**, a016501.
2. Jasin, M. and Rothstein, R. (2013) Repair of strand breaks by homologous recombination. *Cold Spring Harb Perspect Biol*, **5**, a012740.
3. Petronczki, M., Siomos, M.F. and Nasmyth, K. (2003) Un menage a quatre: the molecular biology of chromosome segregation in meiosis. *Cell*, **112**, 423-440.
4. Sarbajna, S. and West, S.C. (2014) Holliday junction processing enzymes as guardians of genome stability. *Trends Biochem Sci*, **39**, 409-419.
5. Mehta, A. and Haber, J.E. (2014) Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb Perspect Biol*, **6**, a016428.
6. Cejka, P., Plank, J.L., Bachrati, C.Z., Hickson, I.D. and Kowalczykowski, S.C. (2010) Rmi1 stimulates decatenation of double Holliday junctions during dissolution by Sgs1-Top3. *Nat Struct Mol Biol*, **17**, 1377-1382.
7. Wu, L. and Hickson, I.D. (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature*, **426**, 870-874.
8. Mankouri, H.W. and Hickson, I.D. (2007) The RecQ helicase-topoisomerase III-Rmi1 complex: a DNA structure-specific 'dissolvasome'? *Trends Biochem Sci*, **32**, 538-546.
9. Wyatt, H.D. and West, S.C. (2014) Holliday junction resolvases. *Cold Spring Harb Perspect Biol*, **6**, a023192.
10. Castor, D., Nair, N., Declais, A.C., Lachaud, C., Toth, R., Macartney, T.J., Lilley, D.M., Arthur, J.S. and Rouse, J. (2013) Cooperative control of holliday junction resolution and DNA repair by the SLX1 and MUS81-EME1 nucleases. *Mol Cell*, **52**, 221-233.
11. Wyatt, H.D., Sarbajna, S., Matos, J. and West, S.C. (2013) Coordinated actions of SLX1-SLX4 and MUS81-EME1 for Holliday junction resolution in human cells. *Mol Cell*, **52**, 234-247.

12. Chen, X.B., Melchionna, R., Denis, C.M., Gaillard, P.H., Blasina, A., Van de Weyer, I., Boddy, M.N., Russell, P., Vialard, J. and McGowan, C.H. (2001) Human Mus81-associated endonuclease cleaves Holliday junctions in vitro. *Mol Cell*, **8**, 1117-1127.
13. Ciccia, A., Constantinou, A. and West, S.C. (2003) Identification and characterization of the human mus81-eme1 endonuclease. *J Biol Chem*, **278**, 25172-25178.
14. Munoz, I.M., Hain, K., Declais, A.C., Gardiner, M., Toh, G.W., Sanchez-Pulido, L., Heuckmann, J.M., Toth, R., Macartney, T., Eppink, B. *et al.* (2009) Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. *Mol Cell*, **35**, 116-127.
15. Svendsen, J.M., Smogorzewska, A., Sowa, M.E., O'Connell, B.C., Gygi, S.P., Elledge, S.J. and Harper, J.W. (2009) Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. *Cell*, **138**, 63-77.
16. Ip, S.C., Rass, U., Blanco, M.G., Flynn, H.R., Skehel, J.M. and West, S.C. (2008) Identification of Holliday junction resolvases from humans and yeast. *Nature*, **456**, 357-361.
17. Rass, U., Compton, S.A., Matos, J., Singleton, M.R., Ip, S.C., Blanco, M.G., Griffith, J.D. and West, S.C. (2010) Mechanism of Holliday junction resolution by the human GEN1 protein. *Genes Dev*, **24**, 1559-1569.
18. Chan, Y.W. and West, S. (2015) GEN1 promotes Holliday junction resolution by a coordinated nick and counter-nick mechanism. *Nucleic Acids Res*, **43**, 10882-10892.
19. Chan, Y.W. and West, S.C. (2014) Spatial control of the GEN1 Holliday junction resolvase ensures genome stability. *Nat Commun*, **5**, 4844.
20. Blanco, M.G., Matos, J. and West, S.C. (2014) Dual control of Yen1 nuclease activity and cellular localization by Cdk and Cdc14 prevents genome instability. *Mol Cell*, **54**, 94-106.
21. Eissler, C.L., Mazon, G., Powers, B.L., Savinov, S.N., Symington, L.S. and Hall, M.C. (2014) The Cdk/cDc14 module controls activation of the Yen1 holliday junction resolvase to promote genome stability. *Mol Cell*, **54**, 80-93.
22. Garcia-Luis, J., Clemente-Blanco, A., Aragon, L. and Machin, F. (2014) Cdc14 targets the Holliday junction resolvase Yen1 to the nucleus in early anaphase. *Cell Cycle*, **13**, 1392-1399.
23. Wild, P. and Matos, J. (2016) Cell cycle control of DNA joint molecule resolution. *Curr Opin Cell Biol*, **40**, 74-80.
24. Matos, J. and West, S.C. (2014) Holliday junction resolution: regulation in space and time. *DNA Repair (Amst)*, **19**, 176-181.
25. Matos, J., Blanco, M.G., Maslen, S., Skehel, J.M. and West, S.C. (2011) Regulatory control of the resolution of DNA recombination intermediates during meiosis and mitosis. *Cell*, **147**, 158-172.
26. Ishikawa, G., Kanai, Y., Takata, K., Takeuchi, R., Shimanouchi, K., Ruike, T., Furukawa, T., Kimura, S. and Sakaguchi, K. (2004) DmGEN, a novel RAD2 family endo-exonuclease from Drosophila melanogaster. *Nucleic Acids Res*, **32**, 6251-6259.
27. Shah Punatar, R., Martin, M.J., Wyatt, H.D., Chan, Y.W. and West, S.C. (2017) Resolution of single and double Holliday junction recombination intermediates by GEN1. *Proc Natl Acad Sci U S A*, **114**, 443-450.

28. Lee, S.H., Princz, L.N., Klugel, M.F., Habermann, B., Pfander, B. and Biertumpfel, C. (2015) Human Holliday junction resolvase GEN1 uses a chromodomain for efficient DNA recognition and cleavage. *Elife*, **4**.
29. Liu, Y., Freeman, A.D., Declais, A.C., Wilson, T.J., Gartner, A. and Lilley, D.M. (2015) Crystal Structure of a Eukaryotic GEN1 Resolving Enzyme Bound to DNA. *Cell Rep*, **13**, 2565-2575.
30. Bellendir, S.P., Rognstad, D.J., Morris, L.P., Zapotoczny, G., Walton, W.G., Redinbo, M.R., Ramsden, D.A., Sekelsky, J. and Erie, D.A. (2017) Substrate preference of Gen endonucleases highlights the importance of branched structures as DNA damage repair intermediates. *Nucleic Acids Res*, **45**, 5333-5348.
31. Kanai, Y., Ishikawa, G., Takeuchi, R., Ruike, T., Nakamura, R., Ihara, A., Ohashi, T., Takata, K., Kimura, S. and Sakaguchi, K. (2007) DmGEN shows a flap endonuclease activity, cleaving the blocked-flap structure and model replication fork. *Febs J*, **274**, 3914-3927.
32. Falquet, B. and Rass, U. (2017) A new role for Holliday junction resolvase Yen1 in processing DNA replication intermediates exposes Dna2 as an accessory replicative helicase. *Microb Cell*, **4**, 32-34.
33. Tsutakawa, S.E., Classen, S., Chapados, B.R., Arvai, A.S., Finger, L.D., Guenther, G., Tomlinson, C.G., Thompson, P., Sarker, A.H., Shen, B. *et al.* (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell*, **145**, 198-211.
34. Patel, N., Atack, J.M., Finger, L.D., Exell, J.C., Thompson, P., Tsutakawa, S., Tainer, J.A., Williams, D.M. and Grasby, J.A. (2012) Flap endonucleases pass 5'-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5'-ends. *Nucleic Acids Res*, **40**, 4507-4519.
35. Orans, J., McSweeney, E.A., Iyer, R.R., Hast, M.A., Hellinga, H.W., Modrich, P. and Beese, L.S. (2011) Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell*, **145**, 212-223.
36. Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, **9**, 671-675.
37. Lilley, D.M. (1985) The kinetic properties of cruciform extrusion are determined by DNA base-sequence. *Nucleic Acids Res*, **13**, 1443-1465.
38. Franke, D., Petoukhov, M.V., Konarev, P.V., Panjkovich, A., Tuukkanen, A., Mertens, H.D.T., Kikhney, A.G., Hajizadeh, N.R., Franklin, J.M., Jeffries, C.M. *et al.* (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr*, **50**, 1212-1225.
39. Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. and Svergun, D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography*, **36**, 1277-1282.
40. Svergun, D.I. (1992) Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *Journal of Applied Crystallography*, **25**, 495-503.
41. Fischer, H., Neto, M.D., Napolitano, H.B., Polikarpov, I. and Craievich, A.F. (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *Journal of Applied Crystallography*, **43**, 101-109.
42. Franke, D. and Svergun, D.I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography*, **42**, 342-346.

43. Volkov, V.V. and Svergun, D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography*, **36**, 860-864.
44. Svergun, D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J*, **76**, 2879-2886.
45. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF chimera - A visualization system for exploratory research and analysis. *J Comput Chem*, **25**, 1605-1612.
46. Svergun, D., Barberato, C. and Koch, M.H.J. (1995) CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography*, **28**, 768-773.
47. Grasby, J.A., Finger, L.D., Tsutakawa, S.E., Atack, J.M. and Tainer, J.A. (2012) Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases. *Trends Biochem Sci*, **37**, 74-84.
48. AlMalki, F.A., Flemming, C.S., Zhang, J., Feng, M., Sedelnikova, S.E., Ceska, T., Rafferty, J.B., Sayers, J.R. and Artymiuk, P.J. (2016) Direct observation of DNA threading in flap endonuclease complexes. *Nat Struct Mol Biol*, **23**, 640-646.
49. Feng, M., Patel, D., Dervan, J.J., Ceska, T., Suck, D., Haq, I. and Sayers, J.R. (2004) Roles of divalent metal ions in flap endonuclease-substrate interactions. *Nat Struct Mol Biol*, **11**, 450-456.
50. Ceska, T.A., Sayers, J.R., Stier, G. and Suck, D. (1996) A helical arch allowing single-stranded DNA to thread through T5 5'-exonuclease. *Nature*, **382**, 90-93.
51. Freeman, A.D., Liu, Y., Declais, A.C., Gartner, A. and Lilley, D.M. (2014) GEN1 from a thermophilic fungus is functionally closely similar to non-eukaryotic junction-resolving enzymes. *J Mol Biol*, **426**, 3946-3959.
52. Mathew, E., Mirza, A. and Menhart, N. (2004) Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins. *J Synchrotron Radiat*, **11**, 314-318.
53. Finger, L.D., Patel, N., Beddows, A., Ma, L., Exell, J.C., Jardine, E., Jones, A.C. and Grasby, J.A. (2013) Observation of unpaired substrate DNA in the flap endonuclease-1 active site. *Nucleic Acids Res*, **41**, 9839-9847.
54. Patel, N., Exell, J.C., Jardine, E., Ombler, B., Finger, L.D., Ciani, B. and Grasby, J.A. (2013) Proline scanning mutagenesis reveals a role for the flap endonuclease-1 helical cap in substrate unpairing. *J Biol Chem*, **288**, 34239-34248.
55. Algasaier, S.I., Exell, J.C., Bennet, I.A., Thompson, M.J., Gotham, V.J., Shaw, S.J., Craggs, T.D., Finger, L.D. and Grasby, J.A. (2016) DNA and Protein Requirements for Substrate Conformational Changes Necessary for Human Flap Endonuclease-1-catalyzed Reaction. *J Biol Chem*, **291**, 8258-8268.
56. Guo, Z., Chavez, V., Singh, P., Finger, L.D., Hang, H., Hegde, M.L. and Shen, B. (2008) Comprehensive mapping of the C-terminus of flap endonuclease-1 reveals distinct interaction sites for five proteins that represent different DNA replication and repair pathways. *J Mol Biol*, **377**, 679-690.

Figures

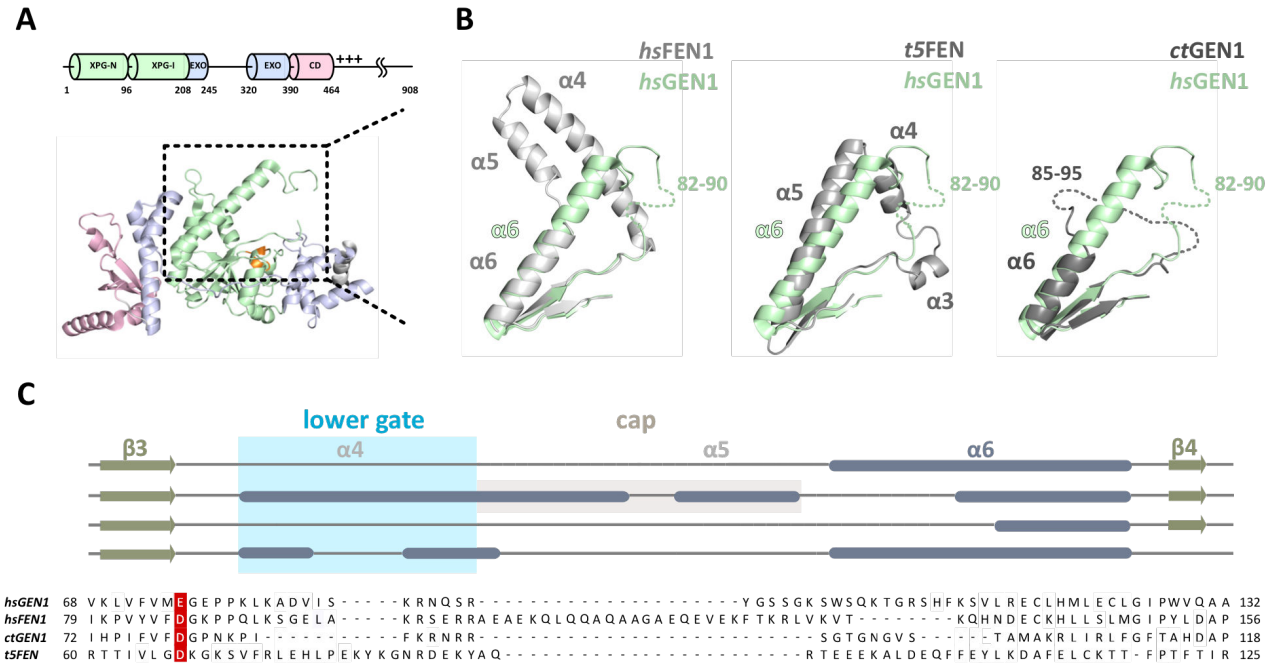


Figure 1. Structure comparison of GEN1 helical arch

(A) The structure of human GEN1 (PDB 5t9j) reveals a conserved catalytic core of Rad2/XPG nucleases. The domain architecture is depicted above the crystal structure. XPG-N and XPG-I domain are in green, 5'-3' exonuclease C-terminal domain (EXO) in blue and chromodomain in pink; a positively charged cluster is trailing the chromodomain. The arch region is highlighted in the black box. (B) Superposition of the helical arches from GEN1-related nucleases. Human GEN1 structure (PDB 5t9j) was superimposed to human FEN1 (PDB 3q8k), bacteriophage T5 FEN (5hml) and *Chaetomium* GEN1 (5co8), respectively. Only the regions corresponding to human GEN1 residues 69-132 were presented. Dotted lines represent the disordered regions in crystal structures. (C) Sequence alignment of the helical arch regions from human GEN1, human FEN1, *Chaetomium* GEN1 and bacteriophage T5 FEN. Secondary structures are presented above the sequences. Dark blue bars represent α helices and green arrows represent β strands. Active site residues are highlighted in red.

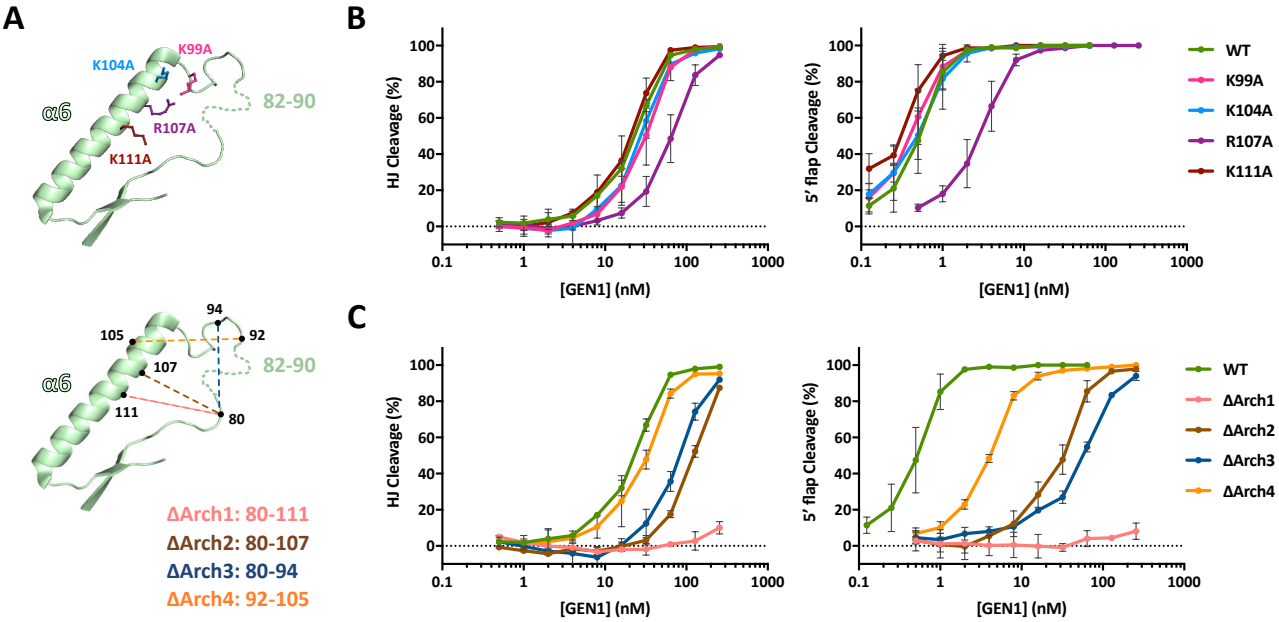


Figure 2. Substrates cleavage of GEN1 point mutations and deletions

(A) Cartoon representation of the arch region of human GEN1 with point mutations used in this study highlighted as sticks in different colors in the upper panel and deletion points indicated in the lower panel. Lys99, Lys104, Arg107 and Lys111 are mutated to alanine to disrupt the positive charge effect on the helical arch. Four arch deletions were generated (Δarch1-4) and the deleted fragments are described below the panel. (B) Substrate cleavage activity of GEN1 on HJs and 5' flaps in comparison to variants with point mutations. (C) Substrate cleavage activity of GEN1 on HJs and 5' flaps in comparison to different arch. 40 nM of 5' 6FAM substrates were mixed with indicated amounts of proteins. Reactions were carried out at 37°C for 15 minutes and loaded on 8% native polyacrylamide gels. Signals were detected by phosphoimager and quantified by ImageJ. The percentage of cleavage was plotted against the enzyme concentration. Each experiment was repeated three times and error bars represent standard deviations.

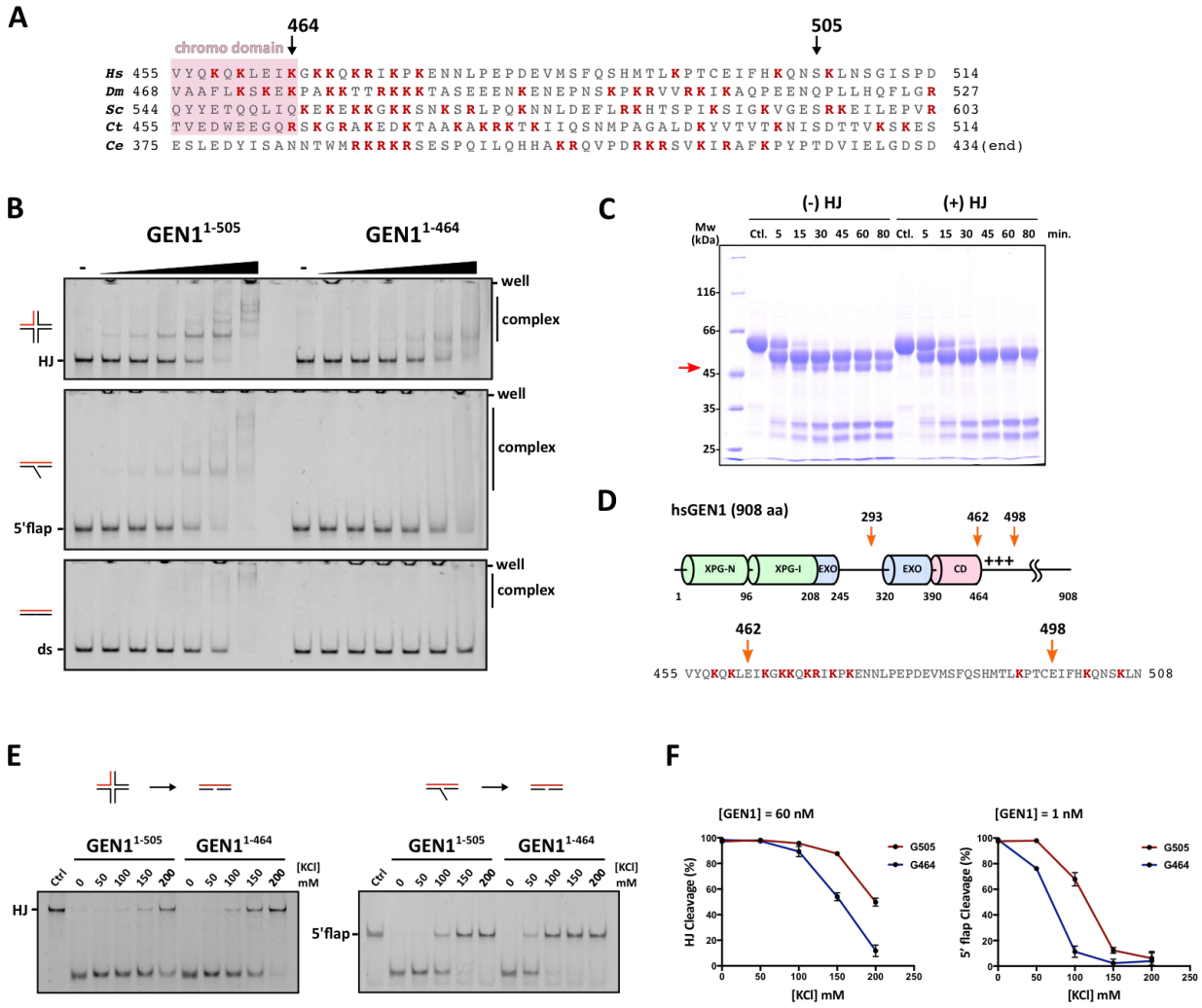
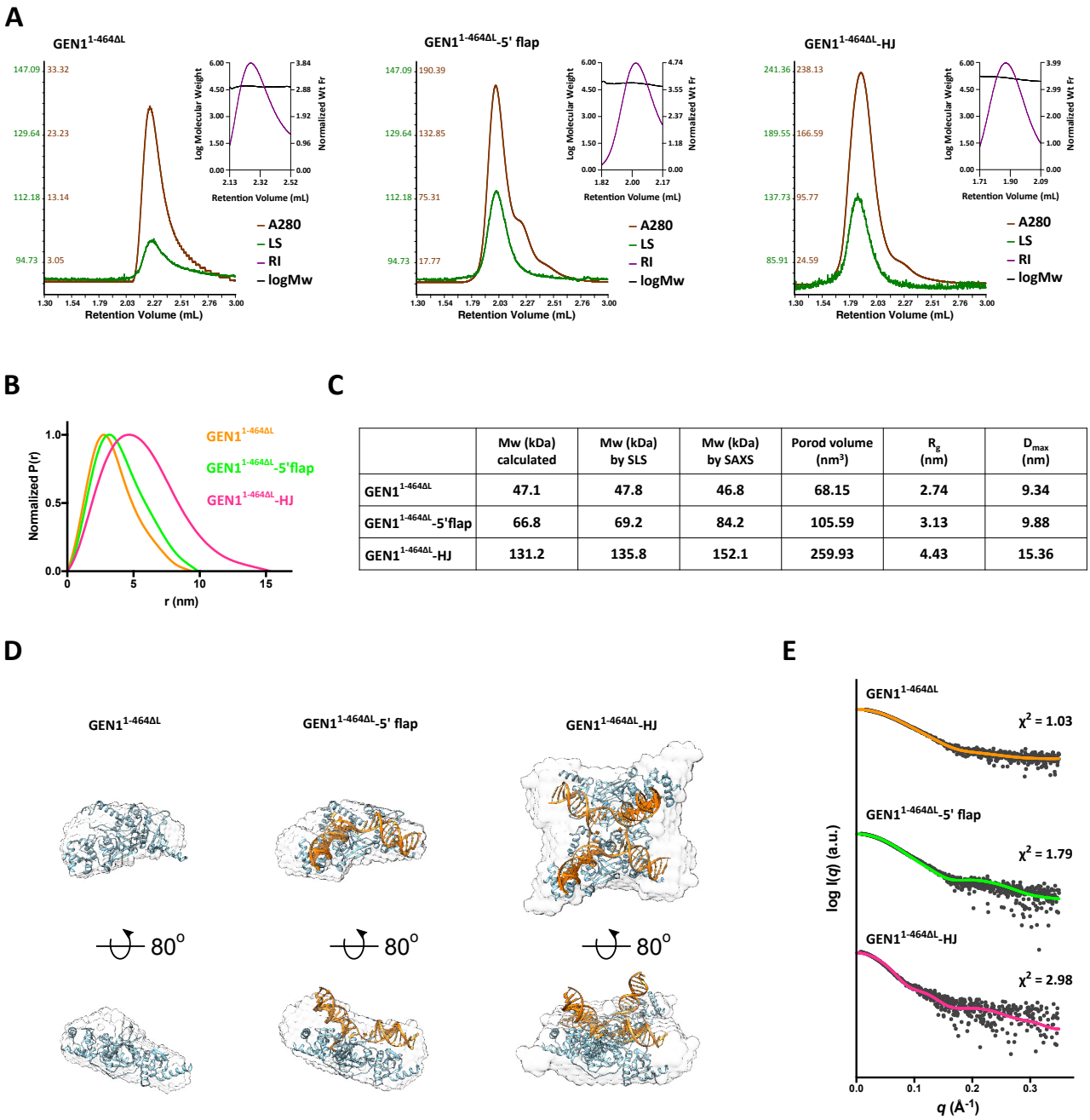


Figure 3. Analysis of the positively charged cluster of GEN1 interacting with DNA

(A) Sequence alignment of positively charged clusters from different GEN1 proteins. The sequences trailing the chromodomain are poorly conserved and predicted as disordered. A polypeptide stretch of 50 residues shows a high propensity of positively charged amino acids. It is noteworthy that *C. elegans* GEN1 lacks a chromodomain, but it also harbors a positively charged cluster at its C-terminus. Basic residues (K/R) are highlighted in red. (B) Electrophoretic mobility shift assays (EMSAs) of GEN1 with different DNA structures. GEN1$^{1\text{-}505}$ and GEN1$^{1\text{-}464}$ (0, 8, 16, 32, 64, 128, 256 nM) were incubated with 40 nM 5' 6FAM labeled Holliday junctions, 5' flap and double stranded DNA in 20 mM Tris-HCl pH 8.0, 50 ng/µl bovine serum albumin, 1 mM dithiothreitol and 5% (v/v) glycerol. Reactions were loaded on 6 % native polyacrylamide gel in 0.5x TBE buffer with additional 1% (v/v) glycerol. Signals were visualized with a phosphoimager. (C) Limited proteolysis of GEN1$^{1\text{-}551\ D30N}$ by GluC protease. 10 µl reactions containing 0.6 mg/ml protein and equal molar amounts of four-way junction (14 bp arm length) were reacted with 3 µl of 0.1 mg/ml GluC protease on ice and stopped with SDS sample buffer at indicated time points. An

additional band appeared in the reactions without four-way junction (indicated with a red arrow). (D) Three major GluC cleavage sites are found for GEN1: C-terminal of Glu293, Glu462 and Glu498. Locations are indicated as orange arrows in the schematic of the domain architecture. Glu293 locates in a non-conserved and disordered region, Glu462 is at the end of the chromodomain, and Glu498 is in the center of the protein and its C-terminal portion is predicted unstructured. The sequence between 462 and 498 consists of a Lys/Arg-rich patch (bottom). Positively charged residues (Lys/Arg) are highlighted in red. (E) Substrates cleavage of $GEN1^{1-505}$ and $GEN1^{1-464}$ on Holliday junctions and 5' flaps in various salt concentrations. 40 nM 5' 6FAM labeled substrates were mixed with 60 nM (for HJs) or 1 nM (for 5' flaps) GEN1 in indicated concentrations of KCl. Reactions were carried out at 37°C for 15 minutes and loaded on 8% native polyacrylamide gels. Fluorescence signals were detected with a phosphoimager. (F) Quantification of salt-dependent cleavage assays by ImageJ. Percentage of cleavage was plotted against KCl concentration. Error bars represent standard deviations of three independent experiments.



	Mw (kDa) calculated	Mw (kDa) by SLS	Mw (kDa) by SAXS	Porod volume (nm^3)	R_g (nm)	D_{max} (nm)
GEN1$^{1\text{-}464\Delta L}$	47.1	47.8	46.8	68.15	2.74	9.34
GEN1$^{1\text{-}464\Delta L}$-5'flap	66.8	69.2	84.2	105.59	3.13	9.88
GEN1$^{1\text{-}464\Delta L}$-HJ	131.2	135.8	152.1	259.93	4.43	15.36

Figure 4. GEN1 dimerizes on Holliday junctions and acts as a monomer on 5' flaps.

(A) Static light –scattering (SLS) analysis of GEN1 with different DNA substrates. The size-exclusion chromatography profiles of GEN1$^{1\text{-}464\Delta L}$ and its complexes with different DNA substrates are presented. The molecular masses were determined from the chosen range of refractive index (highlighted in insets, violet) and mass distributions are depicted in black. (B) The pairwise distance distributions from SAXS experiments for each GEN1$^{1\text{-}464\Delta L}$ sample are shown in different colors. (C) Summary of the determined parameters from SEC-SLS and SEC-SAXS experiments. (D) The solution structures of GEN1$^{1\text{-}464\Delta L}$ alone and in complex with DNA substrates. The *ab initio* models were generated from DAMMIF, averaged by

DAMAVER and refined by DAMMIN. The filtered models are presented as transparent grey envelops. the crystal structure of GEN1$^{1\text{-}505}$ (PDB 5t9j) was fitted into the GEN1$^{1\text{-}464\Delta L}$ envelope. The atomic models of protein-DNA complexes were manually built based on the available GEN1$^{1\text{-}505}$ structure and biochemical data (28). (D) The scattering profile of each GEN1$^{1\text{-}464\Delta L}$ sample. The theoretical scattering profiles of atomic models were fitted to the experimental scattering data and evaluated by the program CRYSOL. The χ^2 values of each fit are given on the side of each scattering curve.

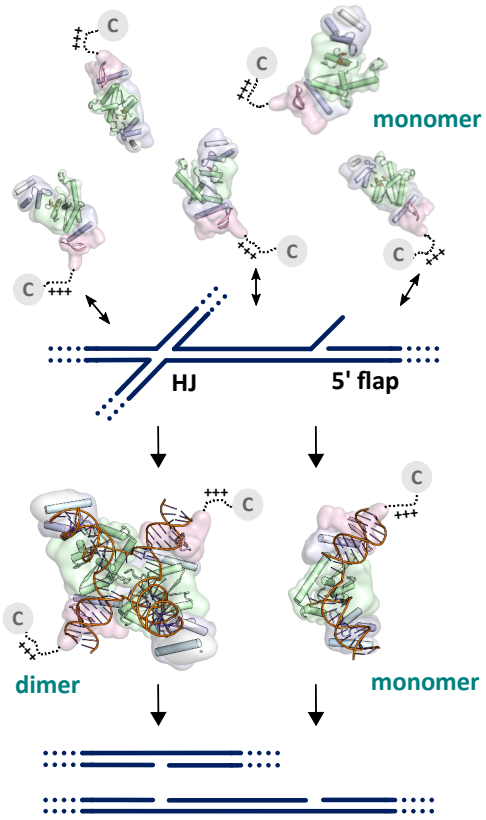


Figure 5. Model of the targeting and discrimination mechanism of GEN1.

GEN1 uses different recognition mechanisms for HJ and 5' flap recognition. Under physiological conditions, GEN1 targets DNA unspecifically via its positively charged cluster, which enhances DNA interactions in general. Once GEN1 is on the DNA, structure-specific interactions select for the correct DNA structures. For Holliday junctions, GEN1 has to recruit a second molecule and assemble as cleavage-competent dimer, which initiates symmetrical incisions and generates ligatable nicked duplexes. For 5' flap structures, the helical arch is required for flap positioning and fast substrate cleavage as a GEN1 monomer.

Supplementary Data

A Oligonucleotides for biochemical assays.

Oligonucleotide	Sequence (5′->3′)
CB209(5′ 6FAM)	ACGCTGCCGAATTCTACCAGTGCCTTGCTAGGACATCTTTGCCCACCTGCAGGTTCACCC
CB210	GGGTGAACCTGCAGGTGGGCAAAGATGTCCATCTGTTGTAATCGTCAAGCTTTATGCCGT
CB211	ACGGCATAAAGCTTGACGATTACAACAGATCATGGAGCTGTCTAGAGGATCCGACTATCG
CB212	CGATAGTCGGATCCTCTAGACAGCTCCATGTAGCAAGGCACTGGTAGAATTCGGCAGCGT
CB215	GGGTGAACCTGCAGGTGGGCAAAGATGTCCTAGCAAGGCACTGGTAGAATTCGGCAGCGT
CB218	GGGTGAACCTGCAGGTGGGCAAAGATGTCC

Four-way junctions were prepared by annealing CB209, CB210, CB211, CB212.
5′ flaps were prepared by annealing CB209, CB212, CB218.
double stranded DNAs were prepared by annealing CB209 and CB215.

B Oligonucleotides for static light scattering and SAXS

Oligonucleotide	Sequence (5′->3′)
4w1515-1	TCGAAGAATTCCGGATTAGGGATGCCGTCT
4w1515-2	AGACGGCATCCCTAAGCTCCATCGTGGCGG
4w1515-3	CCGCCACGATGGAGCCGCTAGGCTCAGAAC
4w1515-4	GTTCTGAGCCTAGCGTCCGGAATTCTTCGA
4w1414-1	CGAAGAATTCCGGATTAGGGATGCCGTC
4w1414-4	TTCTGAGCCTAGCGTCCGGAATTCTTCG
5f14-1	GACGGCATCCCTAA
5f15-1	TCGAAGAATTCCGGA
5f15-2	TTTTGCTCCATCGTGGCGG
5f15-3	CCGCCACGATGGAGCTCCGGAATTCTTCGA

Four-way junctions were prepared by annealing 4w1515-1, 4w1515-2, 4w1515-3, 4w1515-4.
5′ flaps for static light scattering were prepared by annealing 4w1414-1, 4w1414-4, 5f14-1.
5′ flaps for SAXS were prepared by annealing 5f15-1, 5f15-2, 5f15-3.

Table S1. Synthetic oligonucleotides used in this study.

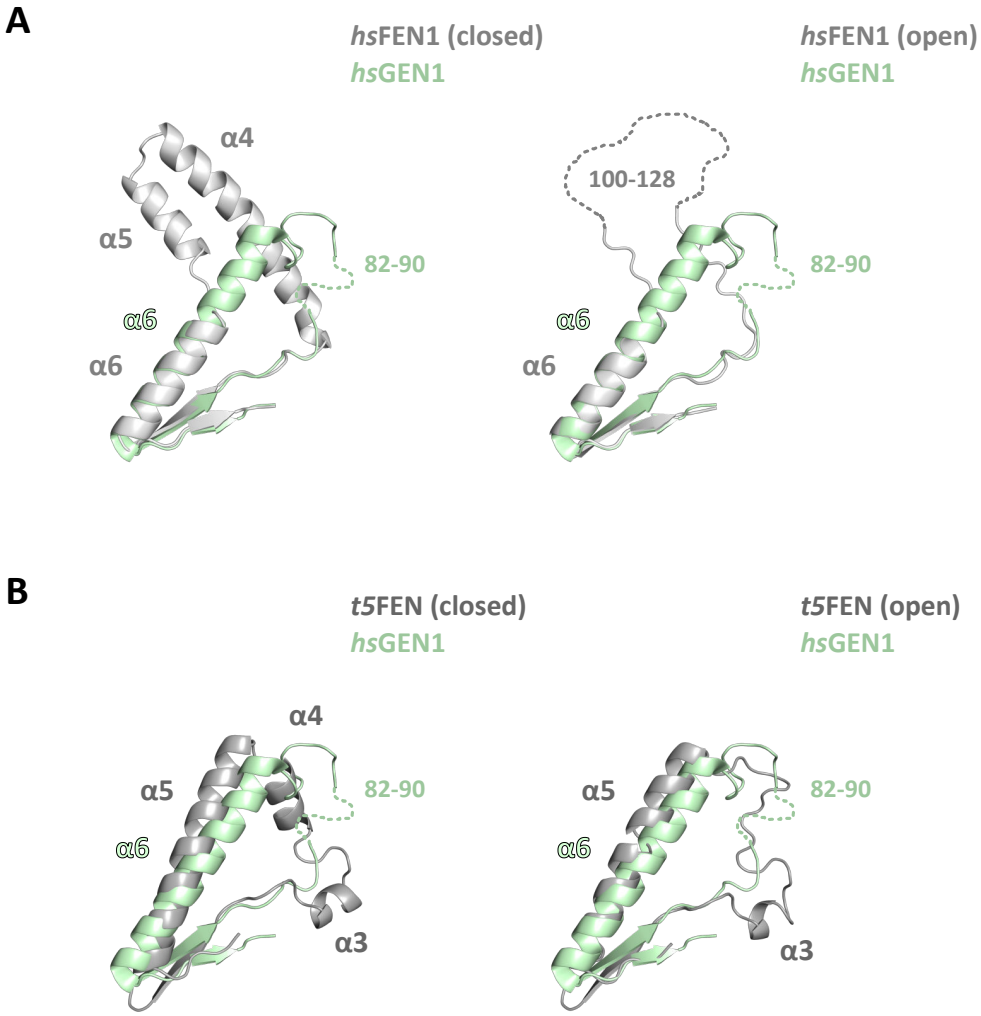


Figure S1. Superposition of the GEN1 helical arch with flap endonucleases in different states.
(A) Superposition of human GEN1 and human FEN1. Crystal structure of human GEN1 (PDB 5t9j) was individually overlaid with human FEN1 in closed- ordered (PDB 3q8k) and open-disordered states (PDB 1ul1). (B) Superposition of human GEN1 with bacteriophage T5 FEN in closed (PDB 5hml, chain A) and open form (PDB 5hml, chain B). Only the arch regions corresponding to human GEN1 residue 69-132 were presented. Dotted lines represent disordered regions in crystal structures and the missing residues are indicated.

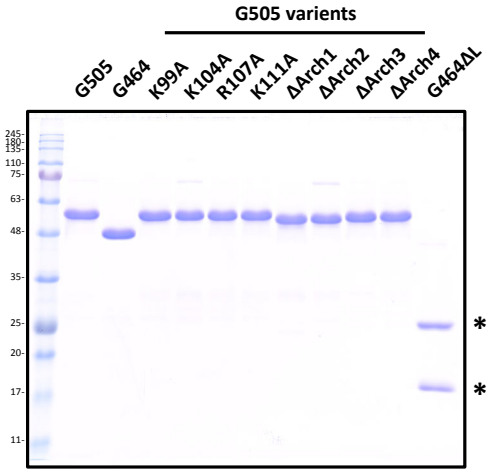


Figure S2. Protein constructs used in this study.

Protein purification procedures are described in Materials and Methods. 2 μg of purified proteins were loaded on 12.5% SDS-PAGE and visualized by Coomassie staining.

GEN1$^{1-505}$

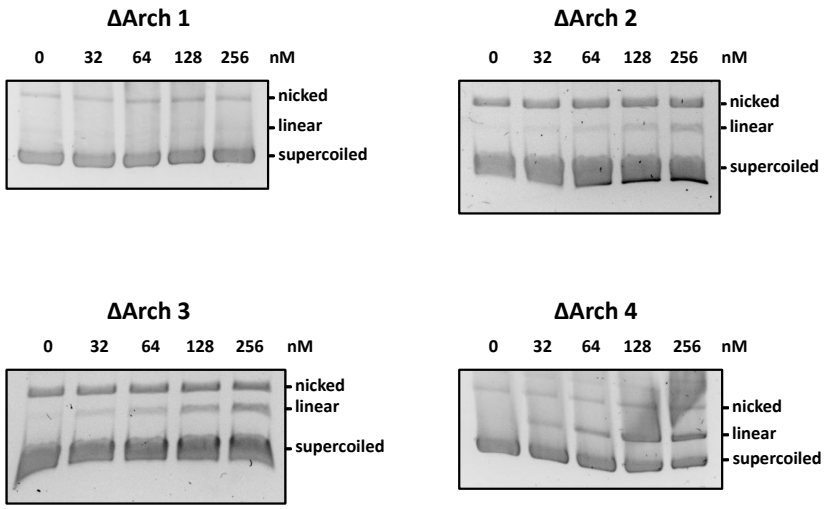


Figure S3. Cruciform-cutting assays of GEN1$^{1-505}$ arch deletions.

Plasmids carrying a cruciform structure were treated with indicated amount of GEN1$^{1-505}$ arch deletions. Only linear products were generated, indicating that the arch deletions do not involved in protein dimerization. Note that the Δarch 1 is catalytic dead, consisting to the result from the assays using synthetic substrates.

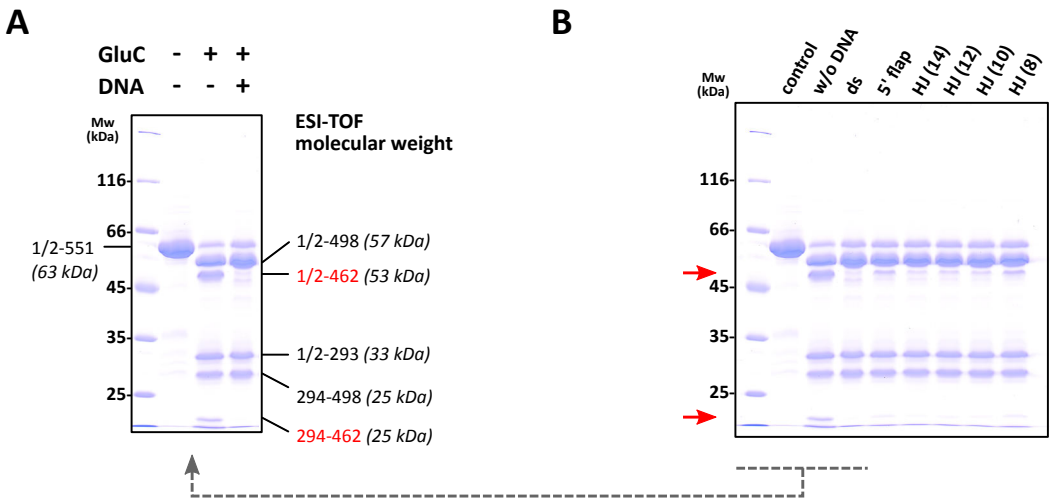


Figure S4. Limited proteolysis of GEN1.

(A) The proteolytic fragments were identified by ESI-TOF mass spectrometry and the coverage as well as approximate molecular weights are indicated. Fragments that are protected from cleavage in the presence of DNA are labeled in red. (B) The protection of proteolysis at Glu462 by different DNA structures and arm lengths indicates that the DNA protection has non-specific in nature. Additional bands in the reaction without DNA are highlighted with red arrows, and the arm lengths of four-way junctions are showed in parentheses.

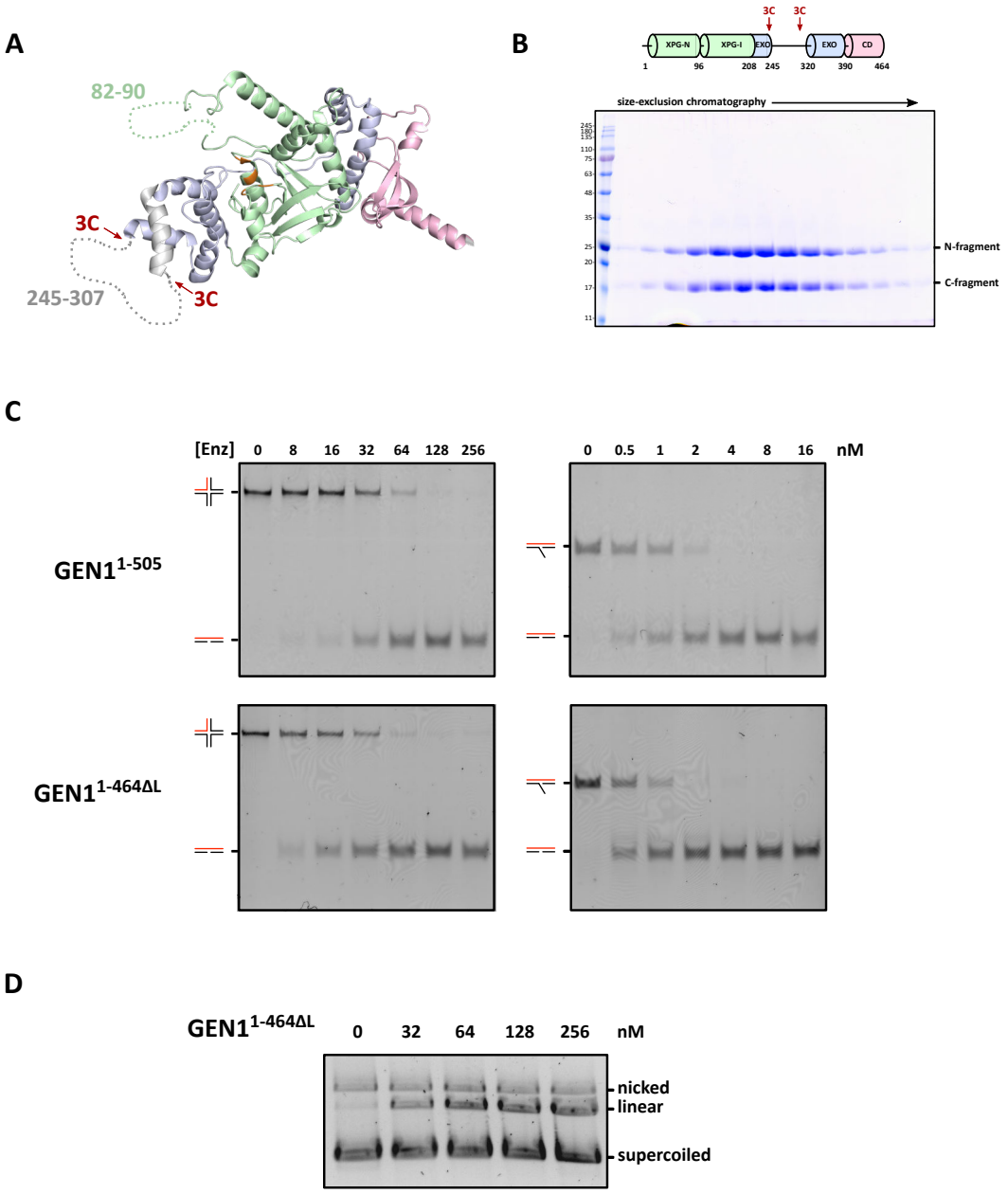


Figure S5. Biochemical characterization of GEN1$^{1\text{-}464\Delta L}$.

(A) Construct design of GEN1$^{1\text{-}464\Delta L}$. HRV 3C protease cleavage sites (LEVLFQ|GP) were introduced after residue 244 and before residue 307, respectively. (B) Domain architecture and purification of GEN1$^{1\text{-}464\Delta L}$. Proteins were treated with 3C protease in order to remove the flexible loop 245-307. Digested fragments associate in size-exclusion chromatography indicating that GEN1$^{1\text{-}464\Delta L}$ remains as a stable nuclease core during purification. (C) Comparison of cleavage activity of GEN1$^{1\text{-}464\Delta L}$ and GEN1$^{1\text{-}505}$. 40 nM 5' 6FAM labeled Holliday junctions and 5' flaps were mixed with indicated amounts of enzyme in 20 mM Tris-HCl pH 8.0, 50 ng/µl BSA, 1 mM DTT and 5 mM $MgCl_2$. GEN1$^{1\text{-}464\Delta L}$ cleaves HJs and 5' flap to the same extent as

GEN1$^{1-505}$. (D) Cruciform plasmid cleavage assay of GEN1$^{1-464\Delta L}$. 50 ng/µl cruciform plasmid pIRbke8^{mut} were mixed with 20 mM Tris-HCl pH 8.0, 50 mM potassium glutamate, 50 ng/µl BSA, 1 mM DTT and 5 mM $MgCl_2$ and incubated at 37°C for 30 minutes. Reactions were initiated by adding indicated amounts of enzyme and incubated at 37°C for 30 minutes. Termination and deproteinization were performed in the same way as for cleavage assays. Products were loaded on 1% 1x TBE native agarose gel and stained with SYBR safe.

Chapter 4 –
Concluding Remarks

4.1. The model of classical HJ resolution in eukaryotes

The work of this thesis presented the first atomic resolution structure of the human HJ resolvase GEN1 that provides the missing structural information for the eukaryotic Rad2/XPG family. The GEN1 structure reveals a strikingly conserved XPG nuclease core with specialized peripheral structure elements for versatile DNA recognition and cleavage features in eukaryotes.

By structural comparison with paralogs, the GEN1-substrate interfaces were identified and further confirmed with biochemical analyses. GEN1 adopted a conserved hydrophobic wedge at the middle of the nuclease core, which inserts into the junction point and separates the upstream and downstream binding interfaces. The DNA arms binding on GEN1 are bended about 90 degrees. Moreover, two structural arrangements enabling the GEN1 dimer to accommodate a four-way junction were identified: (1) an "upper gateway" is created between the upstream interface and the hydrophobic wedge that allows single-stranded DNA connecting to another GEN1 subunit. This gateway is closed in FEN1 by an "acid block" to stabilize the 3′ flap (Tsutakawa et al., 2011). (2) The hydrophobic wedge and the helical arch define a "lower gateway" that allows the cleaving strand passing through. It is noteworthy that the arrangement of the GEN1 helical arch is considerably different from that of human paralogs, FEN1 and EXO1. Instead, the topology of GEN1 helical arch resembles the one from T5 flap endonuclease, a possible explanation is that this architecture is necessary for creating the space for the lower gateway but retains the ability of 5′ flap cleavage for GEN1 as well.

Biochemical studies further highlight the role of the GEN1 helical arch. The mutational analyses indicate that the helical arch plays a role in substrate cleavage. Interestingly, the arch mutations have a bias towards different DNA substrates. Deletion of the arch strongly impairs the cleavage of 5′ flap DNA but shows only a minor effect on HJs, suggesting that this structure is a module for substrate discrimination. The results support that the helical arch is preserved in human GEN1 to enhance the 5′ flap cleavage activity. The specialized

GEN1 homolog in *Chaetomium thermophilum, in* which the helical arch is absent, however, fails to recognize 5′ flaps *in vitro* (Freeman et al., 2014).

The most important finding is that the GEN1 nuclease core is accompanied by a chromodomain, which has not been seen in other members of the Rad2/XPG or any other nuclease family so far. The chromodomain contributes direct interactions to the DNA and has an essential function in substrate recognition and cleavage. Furthermore, our collaboration partners from the Pfander lab have shown that the Yen1 chromodomain is critical for the viability of yeast under DNA damage stress in a *mus81Δ* background. The effect of mutations in the aromatic cage were far more pronounced than a modified chromodomain-DNA interface high-lighting an additional functional role for Yen1's chromodomain. Since chromodomaina have versatile functions in cells such as chromatin-targeting, protein-protein interaction and dimerization, this finding opens many new possibilities for the function role of GEN1 and its chromodomain in maintaining genome integrity.

4.2. Extended discussion

4.2.1. Substrate recognition of GEN1 – Lessons learned from Rad2/XPG nucleases

Eukaryotic XPG nucleases (GEN1, FEN1, EXO1 and XPG in humans) share a conserved catalytic core together with many nucleases from lower organisms, such as T5 FEN (also known as T5 exonuclease), T4 FEN (also known as T4 RNaseH), as well as the 5′ nuclease domain of bacterial DNA polymerase I **(Figure 4-1)** (AlMalki et al., 2016; Eom et al., 1996; Mueser et al., 1996). The common feature of these enzymes is that all of them recognize junction structures. However, the types of recognized junctions are highly diverging from single-stranded nicks, flaps, forks to four-way junctions. Structural studies on these nucleases provide valuable information about their substrate specificity and enable us to have a molecular glance on the divergent evolution of DNA processing. It has been known that the non-conserved linker between XPG-N and XPG-I domains is critical for substrate recognition. The equivalent regions are often disordered in many homologous

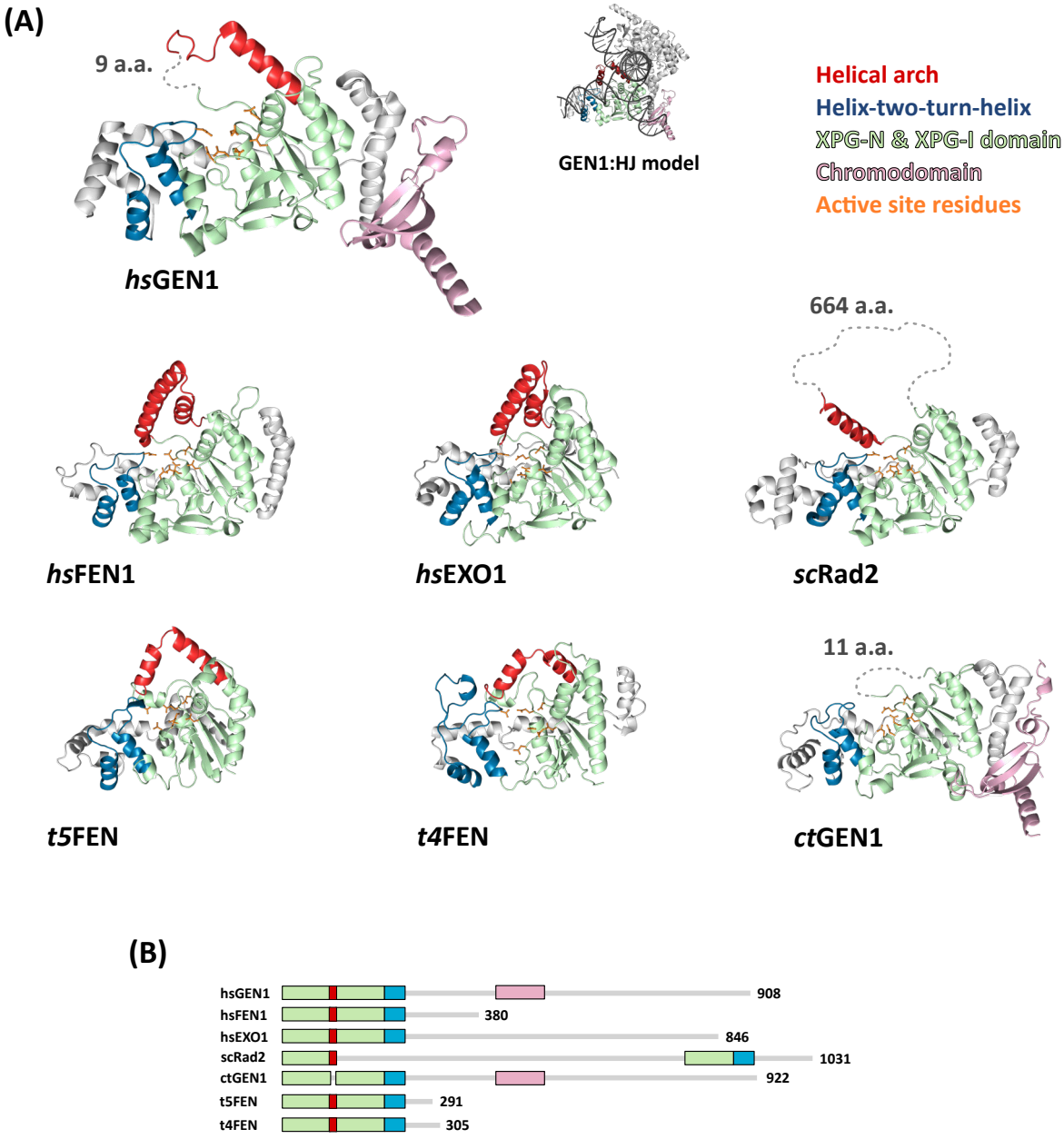


Figure 4-1. Structure comparison of Rad2/XPG nucleases.

(A) Crystal structures of Rad2/XPG nucleases. Critical elements are highlighted in the same colors to compare the topology. Helical arches are in Red, helix-two-turn-helix structures are in dark blue, nuclease cores comprising XPG-N and XPG-I domains are in green, active site residues are in orange, and the chromodomains are in pink. PDB codes: human GEN1 (5T9J), human FEN1 (3Q8K), human EXO1 (3QE9), yeast Rad2 (4Q0W), T5 FEN (5HML), T4 FEN (3H8J), and Chaetomium thermophilum GEN1 (5CO8). (B) Domain architectures of the nucleases. The color representations are the same as above.

proteins, suggesting the linkers are intrinsically dynamic. Interestingly, the linker regions form a stable helical arch or "cap" in the DNA-bound form of human FEN1 and EXO1. This "disorder-to-order" transition of the helical arch facilitates substrate recognition and positioning for efficient nucleolytic cleavage (Orans et al., 2011; Tsutakawa et al., 2011). Similar structural flexibility has been seen in the homologs from lower organisms (AlMalki et al., 2016; Ceska et al., 1996; Grasby et al., 2012). The helical arch is only partially formed in the human GEN1 structure. This is likely due to the missing DNA substrate in the lower gateway. Since the topologic arrangement of the arch region is close to that in T5FEN, human GEN1 most likely uses the similar mechanism to bind and cleave 5' flap substrates. In addition, the T5 FEN-like fold is closer to the arrangement in GEN1 than in human FEN1 as the upper gateway has also an open layout. Therefore, the T5FEN topology could be seen as the solution to GEN1 that is capable of developing the dimerization mechanism for HJ resolution but also preserve the 5' flap cleavage activity.

A "double-nucleotide unpairing mechanism" has been proposed to be the critical step for substrate positioning in FEN1 (Finger et al., 2013; Grasby et al., 2012; Tsutakawa et al., 2011). While the enzyme binds to the substrate, two base pairs of the DNA are unwound from the junction point and enable the scissile phosphate diester to engage with the active site. The similar substrate unpairing movement has also been observed in the GEN1 ortholog from *Chaetomium thermophilum*, suggesting that this mechanism is conserved for the substrate recognition in the GEN1 subfamily (Liu et al., 2015). Unlike other flexible DNA structures such as flaps, bubbles and single-stranded overhangs, Holliday junctions are composed of intact duplex arms and with a rather rigid structure that allows only a limited degree of freedom. Considering the topological constrains of four-way junctions, unwinding two nucleotides on both incised arms causes an axial rotation together with a large conformational change of the junction **(Figure 4-2)**. The structural movement of the four-way junction during GEN1 catalysis is still unknown due to the lack of suitable structural and biophysical information. Perhaps the sequential binding of GEN1 facilitates the junction opening. Nevertheless, the crystal structure presented in this thesis reveals that the GEN1 chromodomain provides additional anchor points to DNA and mostly likely has an important role in stabilizing the unfolded four-way intermediates. It is notable that the GEN1 chromodomain is followed by a positively charged cluster which non-specifically interacts with DNA. Whether this conserved motif has an essential function in facilitating substrate positioning still has to be further investigated.

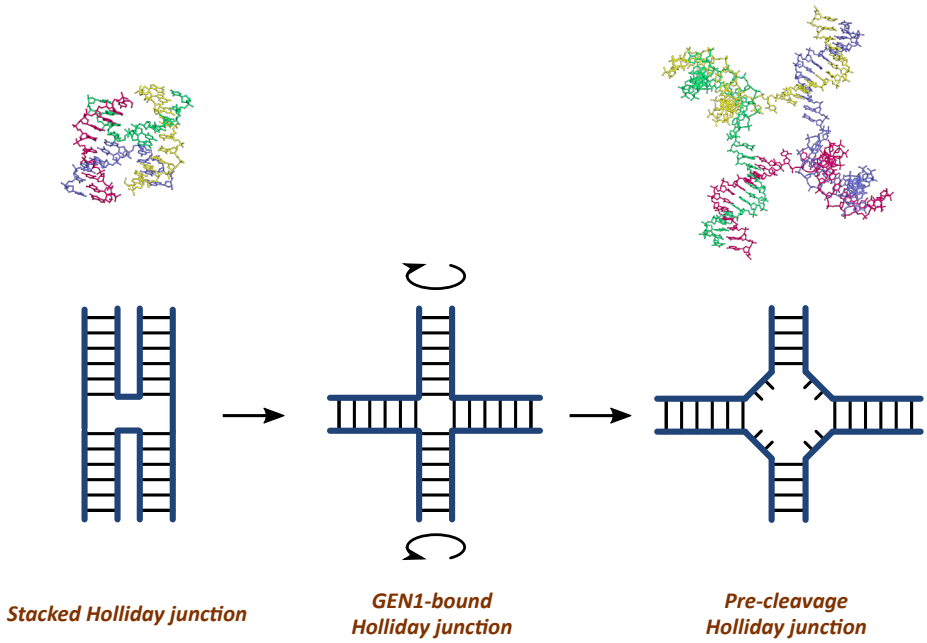


Figure 4-2. The model of the DNA conformational change while GEN1 binding.

In the presence of divalent cations, Holliday junction is stable in a stacked, anti-parallel configuration. While GEN1 binding, the DNA arms are opened and two base pairs from the junction point are disrupted by each GEN1 subunit. DNA axial rotation and large conformational change are necessary to fulfill the double-nucleotide unpairing mechanism. PDB code: stacked HJ (1DCW).

4.2.2. Versatile roles for the GEN1 resolvase

Holliday junctions are the key intermediates during homologous recombination. This structure physically connects two homologous DNA molecules and potentially hampers chromosome segregation. In consequence, it is not surprising that all cells encode enzymes to eliminate these interlocks. Interestingly, resolvases identified in different organisms are rather diverse in sequence but all target the same substrate. The resolution of HJs therefore nicely demonstrates that the convergent evolution of different enzymes can be the solution for a common biological problem.

To faithfully separate HJs, two incisions across the junction are required to generate linear products. Resolvases evolved a dimerization mechanism to assure dual incisions in a precise and coordinated way. The formation of homodimers brings two catalytic centers together at symmetrical positions around the junction point and generates ligatable duplexes after cleavage. Similar characteristics were also found in type II restriction enzymes (Pingoud et al., 2014). All prototypical resolvases in prokaryotes and bacteriophages are stable homodimers in solution, such as *E. coli* RuvC, T4 endonuclease VII and T7 endonuclease I. As an exception, GEN1 is found in its monomeric form in solution and only dimerizes upon binding to four-way junctions. This inducible dimerization likely allows the enzyme to flexibly recognize different types of substrates. Indeed, GEN1 also cleaves 5′ flaps, replication forks and splayed arms *in vitro*. The results in this thesis show the successful reconstitution of human GEN1 in complex with HJs and 5′ flaps using optimized constructs, and it further provides direct evidence that GEN1 interacts with 5′ flaps in monomeric form while it dimerizes on HJs. This observation supports the hypothesis that GEN1 plays versatile roles inside cells. The cleavage of HJs requires the GEN1 dimer formation, probably due to the relative stiffness of four-way junctions, which restrains substrate positioning. As a further consequence, the single-stranded threading mechanism is not possible for a HJ substrate since it does not have a free 5′ end for threading through the helical arch. Therefore, dimerization could be the solution that triggers a conformational change, assembling the cleavage-competent complex which inducing the subsequent substrate remodeling, and feeding of the scissile phosphate diester to the active site. This reflects to the observation that the first cleavage of *Chaetomium* GEN1 is the rate limiting step (Freeman et al., 2014). For those substrates with flexible arms, the monomeric state of GEN1 is enough for the processing by using the same mechanism as T5 flap endonuclease.

The exact biological roles of the 5′ flap endonuclease activity of GEN1 in cells are still unclear. It is possible that GEN1 is a backup system for FEN1, which is the predominant 5′ flap endonuclease to process the intermediates from lagging strand synthesis (Okazaki fragment processing) and long-patch base excision repair (BER). In addition, the asymmetric cleavage of four-way junctions and D-loop structures by MUS81 can also lead to the 5′ flap byproducts.

4.3. Perspectives

Even though the structure of a GEN1-DNA complex has been solved, many questions about GEN1 are still open. In particular, the mechanism of dimerization is a fascinating puzzle to answer. The molecular view on how the enzyme coordinates dual incisions, and the expected conformational changes upon substrate binding have be uncovered by further high-resolution structural studies. One of the difficulties for a biophysical and structural analysis of GEN1-HJ complex is that the protein is highly unstable during complex reconstitution (Lilley, 2017). In this thesis, this bottleneck has been overcome by carefully optimizing the protein constructs. The complex can be examined by biophysical methods and the dimerization feature has been confirmed. Therefore, it is a compelling material for further structural biology and single-molecule studies. In addition, GEN1 is a multi-functional enzyme that cleaves many types of branched DNA structures. However, the roles of cleavage on 5′ flaps, replication forks and splayed arms are poorly investigated. In this thesis, GEN1 constructs with arch deletions were engineered that severely hampered the cleavage of 5′ flaps. Those GEN1 variants could be further applied to cell biology studies to scrutinize the role of the flap endonuclease activity of GEN1 *in vivo*.

It is intriguing to understand how HJ resolution couples to upstream and downstream events during homologous recombination. It has been showed that the RuvC-like activity from cell-free extracts, later identified as GEN1, was co-purified with a branch migration activity (Constantinou et al., 2001). This co-purification was survived after several chromatography separations and the HJ resolution activity is enhanced by the treatment of ATP, strongly suggesting that a helicase subunit is existed and collaborates with GEN1. Moreover, our collaboration partners showed that yeast cells have a higher sensitivity to DNA damaging agents when the point mutations are introduced on the aromatic cage in the Yen1 chromodomain. These results indicate that the chromodomain of GEN1/Yen1 plays additional roles *in vivo*. It is possible that the chromodomain serves as a regulatory module or a docking platform for other endogenous factors. Therefore, further efforts on the identification of GEN1's interaction partners could shed light on the regulation of HJ resolution and GEN1's biological roles.

Abbreviations

6FAM	6-carboxyflurorescein
A. thaliana	*Arabidopsis thaliana*
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluroride hydrochlororide
BER	base excision repair
BIR	break-induced replication
BLM	Bloom's syndrome helicase
BTR	BLM-TOP3α-RMI1-RMI2 complex
C. elegans	*Caenorhabditis elegans*
C. thermophilum	*Chaetomium thermophilum*
CDK1	cyclin-dependent kinase 1
CO	crossover
CtIP	carboxyl terminal binding protein interacting protein
D-loop	displacement loop
D. melanogaster	*Drosophila melanogaster*
dHJ	double Holliday junction
DSB	double-stranded break
dsDNA	double-stranded DNA
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EME1	essential meiotic endonuclease 1
EMSA	electrophoretic mobility shift assay
ERCC1	excision repair cross-complementation group 1
EXO1	exonuclease 1
FEN1	flap endonuclease 1
GEN1	XPG-like endonuclease 1
H2TH	helix-two-turn-helix
HhH	helix-hairpin-helix
HJ	Holliday junction
HR	homologous recombination
IPTG	isopropyl-β-D-thiogalactopyranoside
LOH	loss of heterozygosity
MMR	mismatch repair
Mms4	Methanemethylsulfonate sensitive protein 4
MRN	MRE11-RAD50-NBS1 complex
MRX	Mre11-Rad50-Xrs2 complex
MUS81	MMS and UV sensitive protein 81
NCO	non-crossover
NER	nucleotide excision repair
NES	nuclear export signal
NHEJ	non-homologous end joining
nHJ	nicked Holliday junction
NLS	nuclear localization signal
O. sativa	*Oryza sativa*
PEG	polyethelene glycol
PLK1	Polo-like kinase 1

Rad2	radiation sensitive mutant 2
Rad51	radiation sensitive mutant 51
RMI1	RecQ-mediated genome instability protein 1
RNaseH	ribonuclease H
ROS	reactive oxygen species
RPA	replication protein A
RTEL1	regulator of telomere elongation helicase 1
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharmyces pombe*
Sae2	SUMO-activating enzyme subunit 2
SAXS	small angle X-ray scattering
SCE	sister chromatid exchange
SDSA	synthesis-dependent strand annealing
SEC	size exclusion chromatography
Sgs1	slow growth suppressor 1
SLS	static light scattering
SLX-MUS	SLX1-SLX4-MUS81-EME1 complex
SLX1	synthetic lethal of unknown function 1
SLX4	synthetic lethal of unknown function 4
SMX	SLX1-SLX4-MUS81-EME1-XPF-ERCC1 complex
SSA	single-stranded annealing
ssDNA	single-stranded DNA
SSE	structure-selective endonuclease
STR	Sgs1-Top3-Rmi1 complex
TB	terrific broth
TBE	Tris-borate-EDTA
TCEP	Tris (2-carboxylethyl) phosphine
Top3	topoisomerase III
XPF	xeroderma pigmentosum complementation group F protein
XPG	xeroderma pigmentosum complementation group G protein
XPG-I	XPG internal domain
XPG-N	XPG N-terminal domain
Yen1	crossover junction endodeoxyribonuclease 1

Bibliography

Aboussekhra, A., Chanet, R., Adjiri, A., and Fabre, F. (1992). Semidominant suppressors of Srs2 helicase mutations of Saccharomyces cerevisiae map in the RAD51 gene, whose sequence predicts a protein with similarities to procaryotic RecA proteins. Mol Cell Biol 12, 3224-3234.

Allers, T., and Lichten, M. (2001). Differential timing and control of noncrossover and crossover recombination during meiosis. Cell 106, 47-57.

AlMalki, F.A., Flemming, C.S., Zhang, J., Feng, M., Sedelnikova, S.E., Ceska, T., Rafferty, J.B., Sayers, J.R., and Artymiuk, P.J. (2016). Direct observation of DNA threading in flap endonuclease complexes. Nat Struct Mol Biol 23, 640-646.

Anand, R., Ranjha, L., Cannavo, E., and Cejka, P. (2016). Phosphorylated CtIP Functions as a Co-factor of the MRE11-RAD50-NBS1 Endonuclease in DNA End Resection. Mol Cell 64, 940-950.

Andersen, S.L., Bergstralh, D.T., Kohl, K.P., LaRocque, J.R., Moore, C.B., and Sekelsky, J. (2009). Drosophila MUS312 and the vertebrate ortholog BTBD12 interact with DNA structure-specific endonucleases in DNA repair and recombination. Mol Cell 35, 128-135.

Andersen, S.L., Kuo, H.K., Savukoski, D., Brodsky, M.H., and Sekelsky, J. (2011). Three structure-selective endonucleases are essential in the absence of BLM helicase in Drosophila. PLoS Genet 7, e1002315.

Bailly, A.P., Freeman, A., Hall, J., Declais, A.C., Alpi, A., Lilley, D.M., Ahmed, S., and Gartner, A. (2010). The Caenorhabditis elegans homolog of Gen1/Yen1 resolvases links DNA damage signaling to DNA double-strand break repair. PLoS Genet 6, e1001025.

Barber, L.J., Youds, J.L., Ward, J.D., McIlwraith, M.J., O'Neil, N.J., Petalcorin, M.I., Martin, J.S., Collis, S.J., Cantor, S.B., Auclair, M., et al. (2008). RTEL1 maintains genomic stability by suppressing homologous recombination. Cell 135, 261-271.

Basile, G., Aker, M., and Mortimer, R.K. (1992). Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene RAD51. Mol Cell Biol 12, 3235-3246.

Bauknecht, M., and Kobbe, D. (2014). AtGEN1 and AtSEND1, two paralogs in Arabidopsis, possess holliday junction resolvase activity. Plant Physiol 166, 202-216.

Bellendir, S.P., Rognstad, D.J., Morris, L.P., Zapotoczny, G., Walton, W.G., Redinbo, M.R., Ramsden, D.A., Sekelsky, J., and Erie, D.A. (2017). Substrate preference of Gen endonucleases highlights the importance of branched structures as DNA damage repair intermediates. Nucleic Acids Res 45, 5333-5348.

Bennett, R.J., Dunderdale, H.J., and West, S.C. (1993). Resolution of Holliday junctions by RuvC resolvase: cleavage specificity and DNA distortion. Cell 74, 1021-1031.

Bergerat, A., de Massy, B., Gadelle, D., Varoutas, P.C., Nicolas, A., and Forterre, P. (1997). An atypical topoisomerase II from Archaea with implications for meiotic recombination. Nature 386, 414-417.

Bianco, P.R., Tracy, R.B., and Kowalczykowski, S.C. (1998). DNA strand exchange proteins: a biochemical and physical comparison. Front Biosci 3, D570-603.

Bizard, A.H., and Hickson, I.D. (2014). The dissolution of double Holliday junctions. Cold Spring Harb Perspect Biol 6, a016477.

Blackwood, J.K., Rzechorzek, N.J., Bray, S.M., Maman, J.D., Pellegrini, L., and Robinson, N.P. (2013). End-resection at DNA double-strand breaks in the three domains of life. Biochem Soc Trans 41, 314-320.

Blanco, M.G., and Matos, J. (2015). Hold your horSSEs: controlling structure-selective endonucleases MUS81 and Yen1/GEN1. Front Genet 6, 253.

Blanco, M.G., Matos, J., Rass, U., Ip, S.C., and West, S.C. (2010). Functional overlap between the structure-specific nucleases Yen1 and Mus81-Mms4 for DNA-damage repair in S. cerevisiae. DNA Repair (Amst) 9, 394-402.

Blanco, M.G., Matos, J., and West, S.C. (2014). Dual control of Yen1 nuclease activity and cellular localization by Cdk and Cdc14 prevents genome instability. Mol Cell 54, 94-106.

Bocquet, N., Bizard, A.H., Abdulrahman, W., Larsen, N.B., Faty, M., Cavadini, S., Bunker, R.D., Kowalczykowski, S.C., Cejka, P., Hickson, I.D., et al. (2014). Structural and mechanistic insight into Holliday-junction dissolution by topoisomerase IIIalpha and RMI1. Nat Struct Mol Biol 21, 261-268.

Boddy, M.N., Gaillard, P.H., McDonald, W.H., Shanahan, P., Yates, J.R., 3rd, and Russell, P. (2001). Mus81-Eme1 are essential components of a Holliday junction resolvase. Cell 107, 537-548.

Bzymek, M., Thayer, N.H., Oh, S.D., Kleckner, N., and Hunter, N. (2010). Double Holliday junctions are intermediates of DNA break repair. Nature 464, 937-941.

Cannavo, E., and Cejka, P. (2014). Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. Nature 514, 122-125.

Cannavo, E., Cejka, P., and Kowalczykowski, S.C. (2013). Relationship of DNA degradation by Saccharomyces cerevisiae exonuclease 1 and its stimulation by RPA and Mre11-Rad50-Xrs2 to DNA end resection. Proc Natl Acad Sci U S A 110, E1661-1668.

Castor, D., Nair, N., Declais, A.C., Lachaud, C., Toth, R., Macartney, T.J., Lilley, D.M., Arthur, J.S., and Rouse, J. (2013). Cooperative control of holliday junction resolution and DNA repair by the SLX1 and MUS81-EME1 nucleases. Mol Cell 52, 221-233.

Cejka, P., Cannavo, E., Polaczek, P., Masuda-Sasa, T., Pokharel, S., Campbell, J.L., and Kowalczykowski, S.C. (2010a). DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2. Nature 467, 112-116.

Cejka, P., Plank, J.L., Bachrati, C.Z., Hickson, I.D., and Kowalczykowski, S.C. (2010b). Rmi1 stimulates decatenation of double Holliday junctions during dissolution by Sgs1-Top3. Nat Struct Mol Biol 17, 1377-1382.

Cejka, P., Plank, J.L., Dombrowski, C.C., and Kowalczykowski, S.C. (2012). Decatenation of DNA by the S. cerevisiae Sgs1-Top3-Rmi1 and RPA complex: a mechanism for disentangling chromosomes. Mol Cell 47, 886-896.

Ceska, T.A., Sayers, J.R., Stier, G., and Suck, D. (1996). A helical arch allowing single-stranded DNA to thread through T5 5'-exonuclease. Nature 382, 90-93.

Chaganti, R.S., Schonberg, S., and German, J. (1974). A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes. Proc Natl Acad Sci U S A 71, 4508-4512.

Chan, Y.W., and West, S. (2015). GEN1 promotes Holliday junction resolution by a coordinated nick and counter-nick mechanism. Nucleic Acids Res 43, 10882-10892.

Chan, Y.W., and West, S.C. (2014). Spatial control of the GEN1 Holliday junction resolvase ensures genome stability. Nat Commun 5, 4844.

Chang, M., Bellaoui, M., Zhang, C., Desai, R., Morozov, P., Delgado-Cruzata, L., Rothstein, R., Freyer, G.A., Boone, C., and Brown, G.W. (2005). RMI1/NCE4, a suppressor of genome instability, encodes a member of the RecQ helicase/Topo III complex. EMBO J 24, 2024-2033.

Chen, X.B., Melchionna, R., Denis, C.M., Gaillard, P.H., Blasina, A., Van de Weyer, I., Boddy, M.N., Russell, P., Vialard, J., and McGowan, C.H. (2001). Human Mus81-associated endonuclease cleaves Holliday junctions in vitro. Mol Cell 8, 1117-1127.

Chi, P., San Filippo, J., Sehorn, M.G., Petukhova, G.V., and Sung, P. (2007). Bipartite stimulatory action of the Hop2-Mnd1 complex on the Rad51 recombinase. Genes Dev 21, 1747-1757.

Ciccia, A., Constantinou, A., and West, S.C. (2003). Identification and characterization of the human mus81-eme1 endonuclease. J Biol Chem 278, 25172-25178.

Ciccia, A., McDonald, N., and West, S.C. (2008). Structural and functional relationships of the XPF/MUS81 family of proteins. Annu Rev Biochem 77, 259-287.

Constantinou, A., Chen, X.B., McGowan, C.H., and West, S.C. (2002). Holliday junction resolution in human cells: two junction endonucleases with distinct substrate specificities. EMBO J 21, 5577-5585.

Constantinou, A., Davies, A.A., and West, S.C. (2001). Branch migration and Holliday junction resolution catalyzed by activities from mammalian cells. Cell 104, 259-268.

Cromie, G.A., Hyppa, R.W., Taylor, A.F., Zakharyevich, K., Hunter, N., and Smith, G.R. (2006). Single Holliday junctions are intermediates of meiotic recombination. Cell 127, 1167-1178.

Cybulski, K.E., and Howlett, N.G. (2011). FANCP/SLX4: a Swiss army knife of DNA interstrand crosslink repair. Cell Cycle 10, 1757-1763.

Daley, J.M., and Sung, P. (2014). 53BP1, BRCA1, and the choice between recombination and end joining at DNA double-strand breaks. Mol Cell Biol 34, 1380-1388.

Dehe, P.M., and Gaillard, P.H. (2017). Control of structure-specific endonucleases to maintain genome stability. Nat Rev Mol Cell Biol 18, 315-330.

Eissler, C.L., Mazon, G., Powers, B.L., Savinov, S.N., Symington, L.S., and Hall, M.C. (2014). The Cdk/cDc14 module controls activation of the Yen1 holliday junction resolvase to promote genome stability. Mol Cell 54, 80-93.

Elborough, K.M., and West, S.C. (1990). Resolution of synthetic Holliday junctions in DNA by an endonuclease activity from calf thymus. EMBO J 9, 2931-2936.

Eom, S.H., Wang, J., and Steitz, T.A. (1996). Structure of Taq polymerase with DNA at the polymerase active site. Nature 382, 278-281.

Fekairi, S., Scaglione, S., Chahwan, C., Taylor, E.R., Tissier, A., Coulon, S., Dong, M.Q., Ruse, C., Yates, J.R., 3rd, Russell, P., et al. (2009). Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. Cell 138, 78-89.

Ferguson, D.O., and Holloman, W.K. (1996). Recombinational repair of gaps in DNA is asymmetric in Ustilago maydis and can be explained by a migrating D-loop model. Proc Natl Acad Sci U S A 93, 5419-5424.

Finger, L.D., Patel, N., Beddows, A., Ma, L., Exell, J.C., Jardine, E., Jones, A.C., and Grasby, J.A. (2013). Observation of unpaired substrate DNA in the flap endonuclease-1 active site. Nucleic Acids Res 41, 9839-9847.

Freeman, A.D., Liu, Y., Declais, A.C., Gartner, A., and Lilley, D.M. (2014). GEN1 from a thermophilic fungus is functionally closely similar to non-eukaryotic junction-resolving enzymes. J Mol Biol 426, 3946-3959.

Fricke, W.M., and Brill, S.J. (2003). Slx1-Slx4 is a second structure-specific endonuclease functionally redundant with Sgs1-Top3. Genes Dev 17, 1768-1778.

Furuse, M., Nagase, Y., Tsubouchi, H., Murakami-Murofushi, K., Shibata, T., and Ohta, K. (1998). Distinct roles of two separable in vitro activities of yeast Mre11 in mitotic and meiotic recombination. EMBO J 17, 6412-6425.

Gallo-Fernandez, M., Saugar, I., Ortiz-Bazan, M.A., Vazquez, M.V., and Tercero, J.A. (2012). Cell cycle-dependent regulation of the nuclease activity of Mus81-Eme1/Mms4. Nucleic Acids Res 40, 8325-8335.

Gangloff, S., McDonald, J.P., Bendixen, C., Arthur, L., and Rothstein, R. (1994). The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. Mol Cell Biol 14, 8391-8398.

Garcia-Luis, J., Clemente-Blanco, A., Aragon, L., and Machin, F. (2014). Cdc14 targets the Holliday junction resolvase Yen1 to the nucleus in early anaphase. Cell Cycle 13, 1392-1399.

Garner, E., Kim, Y., Lach, F.P., Kottemann, M.C., and Smogorzewska, A. (2013). Human GEN1 and the SLX4-associated nucleases MUS81 and SLX1 are essential for the resolution of replication-induced Holliday junctions. Cell Rep 5, 207-215.

Grasby, J.A., Finger, L.D., Tsutakawa, S.E., Atack, J.M., and Tainer, J.A. (2012). Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases. Trends Biochem Sci 37, 74-84.

Hanada, K., Budzowska, M., Davies, S.L., van Drunen, E., Onizawa, H., Beverloo, H.B., Maas, A., Essers, J., Hickson, I.D., and Kanaar, R. (2007). The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks. Nat Struct Mol Biol 14, 1096-1104.

Hartsuiker, E., Neale, M.J., and Carr, A.M. (2009). Distinct requirements for the Rad32(Mre11) nuclease and Ctp1(CtIP) in the removal of covalently bound topoisomerase I and II from DNA. Mol Cell 33, 117-123.

Heyer, W.D. (2015). Regulation of recombination and genomic maintenance. Cold Spring Harb Perspect Biol 7, a016501.

Holliday, R. (1964). The Induction of Mitotic Recombination by Mitomycin C in Ustilago and Saccharomyces. Genetics 50, 323-335.

Hollingsworth, N.M., and Brill, S.J. (2004). The Mus81 solution to resolution: generating meiotic crossovers without Holliday junctions. Genes Dev 18, 117-125.

Hyde, H., Davies, A.A., Benson, F.E., and West, S.C. (1994). Resolution of recombination intermediates by a mammalian activity functionally analogous to Escherichia coli RuvC resolvase. J Biol Chem 269, 5202-5209.

Ip, S.C., Rass, U., Blanco, M.G., Flynn, H.R., Skehel, J.M., and West, S.C. (2008). Identification of Holliday junction resolvases from humans and yeast. Nature 456, 357-361.

Ira, G., Malkova, A., Liberi, G., Foiani, M., and Haber, J.E. (2003). Srs2 and Sgs1-Top3 suppress crossovers during double-strand break repair in yeast. Cell 115, 401-411.

Ira, G., Pellicioli, A., Balijja, A., Wang, X., Fiorani, S., Carotenuto, W., Liberi, G., Bressan, D., Wan, L., Hollingsworth, N.M., et al. (2004). DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. Nature 431, 1011-1017.

Ishikawa, G., Kanai, Y., Takata, K., Takeuchi, R., Shimanouchi, K., Ruike, T., Furukawa, T., Kimura, S., and Sakaguchi, K. (2004). DmGEN, a novel RAD2 family endo-exonuclease from Drosophila melanogaster. Nucleic Acids Res 32, 6251-6259.

Jasin, M., and Rothstein, R. (2013). Repair of strand breaks by homologous recombination. Cold Spring Harb Perspect Biol 5, a012740.

Jensen, R.B., Carreira, A., and Kowalczykowski, S.C. (2010). Purified human BRCA2 stimulates RAD51-mediated recombination. Nature 467, 678-683.

Kanai, Y., Ishikawa, G., Takeuchi, R., Ruike, T., Nakamura, R., Ihara, A., Ohashi, T., Takata, K., Kimura, S., and Sakaguchi, K. (2007). DmGEN shows a flap endonuclease activity, cleaving the blocked-flap structure and model replication fork. FEBS J 274, 3914-3927.

Keeney, S., Giroux, C.N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell 88, 375-384.

Kowalczykowski, S.C. (2015). An Overview of the Molecular Mechanisms of Recombinational DNA Repair. Cold Spring Harb Perspect Biol 7.

Kuzminov, A. (2001). Single-strand interruptions in replicating chromosomes cause double-strand breaks. Proc Natl Acad Sci U S A 98, 8241-8246.

Lee, J.H., Ghirlando, R., Bhaskara, V., Hoffmeyer, M.R., Gu, J., and Paull, T.T. (2003). Regulation of Mre11/Rad50 by Nbs1: effects on nucleotide-dependent DNA binding and association with ataxia-telangiectasia-like disorder mutant complexes. J Biol Chem 278, 45171-45181.

Li, W., and Wang, J.C. (1998). Mammalian DNA topoisomerase IIIalpha is essential in early embryogenesis. Proc Natl Acad Sci U S A 95, 1010-1013.

Lilley, D.M., and Markham, A.F. (1983). Dynamics of cruciform extrusion in supercoiled DNA: use of a synthetic inverted repeat to study conformational populations. EMBO J 2, 527-533.

Lilley, D.M.J. (2017). Holliday junction-resolving enzymes-structures and mechanisms. FEBS Lett 591, 1073-1082.

Liu, Y., Freeman, A.D., Declais, A.C., Wilson, T.J., Gartner, A., and Lilley, D.M. (2015). Crystal Structure of a Eukaryotic GEN1 Resolving Enzyme Bound to DNA. Cell Rep 13, 2565-2575.

Lorenz, A., West, S.C., and Whitby, M.C. (2010). The human Holliday junction resolvase GEN1 rescues the meiotic phenotype of a Schizosaccharomyces pombe mus81 mutant. Nucleic Acids Res 38, 1866-1873.

Malkova, A., Ivanov, E.L., and Haber, J.E. (1996). Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. Proc Natl Acad Sci U S A 93, 7131-7136.

Matos, J., Blanco, M.G., Maslen, S., Skehel, J.M., and West, S.C. (2011). Regulatory control of the resolution of DNA recombination intermediates during meiosis and mitosis. Cell 147, 158-172.

Matos, J., and West, S.C. (2014). Holliday junction resolution: regulation in space and time. DNA Repair (Amst) 19, 176-181.

Mehta, A., and Haber, J.E. (2014). Sources of DNA double-strand breaks and models of recombinational DNA repair. Cold Spring Harb Perspect Biol 6, a016428.

Mimitou, E.P., and Symington, L.S. (2008). Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature 455, 770-774.

Mimitou, E.P., and Symington, L.S. (2009). Nucleases and helicases take center stage in homologous recombination. Trends Biochem Sci 34, 264-272.

Mimitou, E.P., and Symington, L.S. (2011). DNA end resection--unraveling the tail. DNA Repair (Amst) 10, 344-348.

Mueser, T.C., Nossal, N.G., and Hyde, C.C. (1996). Structure of bacteriophage T4 RNase H, a 5' to 3' RNA-DNA and DNA-DNA exonuclease with sequence similarity to the RAD2 family of eukaryotic proteins. Cell 85, 1101-1112.

Mullen, J.R., Kaliraman, V., Ibrahim, S.S., and Brill, S.J. (2001). Requirement for three novel protein complexes in the absence of the Sgs1 DNA helicase in Saccharomyces cerevisiae. Genetics 157, 103-118.

Mullen, J.R., Nallaseth, F.S., Lan, Y.Q., Slagle, C.E., and Brill, S.J. (2005). Yeast Rmi1/Nce4 controls genome stability as a subunit of the Sgs1-Top3 complex. Mol Cell Biol 25, 4476-4487.

Munoz, I.M., Hain, K., Declais, A.C., Gardiner, M., Toh, G.W., Sanchez-Pulido, L., Heuckmann, J.M., Toth, R., Macartney, T., Eppink, B., et al. (2009). Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. Mol Cell 35, 116-127.

Nassif, N., Penney, J., Pal, S., Engels, W.R., and Gloor, G.B. (1994). Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. Mol Cell Biol 14, 1613-1625.

Nimonkar, A.V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J.L., Wyman, C., Modrich, P., and Kowalczykowski, S.C. (2011). BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. Genes Dev 25, 350-362.

Nimonkar, A.V., Ozsoy, A.Z., Genschel, J., Modrich, P., and Kowalczykowski, S.C. (2008). Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair. Proc Natl Acad Sci U S A 105, 16906-16911.

Nishino, T., Ishino, Y., and Morikawa, K. (2006). Structure-specific DNA nucleases: structural basis for 3D-scissors. Curr Opin Struct Biol 16, 60-67.

Niu, H., Chung, W.H., Zhu, Z., Kwon, Y., Zhao, W., Chi, P., Prakash, R., Seong, C., Liu, D., Lu, L., et al. (2010). Mechanism of the ATP-dependent DNA end-resection machinery from Saccharomyces cerevisiae. Nature 467, 108-111.

Ogawa, T., Yu, X., Shinohara, A., and Egelman, E.H. (1993). Similarity of the yeast RAD51 filament to the bacterial RecA filament. Science 259, 1896-1899.

Orans, J., McSweeney, E.A., Iyer, R.R., Hast, M.A., Hellinga, H.W., Modrich, P., and Beese, L.S. (2011). Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. Cell 145, 212-223.

Osman, F., Dixon, J., Doe, C.L., and Whitby, M.C. (2003). Generating crossovers by resolution of nicked Holliday junctions: a role for Mus81-Eme1 in meiosis. Mol Cell 12, 761-774.

Petermann, E., and Helleday, T. (2010). Pathways of mammalian replication fork restart. Nat Rev Mol Cell Biol 11, 683-687.

Pezza, R.J., Voloshin, O.N., Vanevski, F., and Camerini-Otero, R.D. (2007). Hop2/Mnd1 acts on two critical steps in Dmc1-promoted homologous pairing. Genes Dev 21, 1758-1766.

Pingoud, A., Wilson, G.G., and Wende, W. (2014). Type II restriction endonucleases--a historical perspective and more. Nucleic Acids Res 42, 7489-7527.

Plank, J.L., Wu, J., and Hsieh, T.S. (2006). Topoisomerase IIIalpha and Bloom's helicase can resolve a mobile double Holliday junction substrate through convergent branch migration. Proc Natl Acad Sci U S A 103, 11118-11123.

Postow, L., Ullsperger, C., Keller, R.W., Bustamante, C., Vologodskii, A.V., and Cozzarelli, N.R. (2001). Positive torsional strain causes the formation of a four-way junction at replication forks. J Biol Chem 276, 2790-2796.

Rass, U., Compton, S.A., Matos, J., Singleton, M.R., Ip, S.C., Blanco, M.G., Griffith, J.D., and West, S.C. (2010). Mechanism of Holliday junction resolution by the human GEN1 protein. Genes Dev 24, 1559-1569.

Resnick, M.A. (1976). The repair of double-strand breaks in DNA; a model involving recombination. J Theor Biol 59, 97-106.

San Filippo, J., Chi, P., Sehorn, M.G., Etchin, J., Krejci, L., and Sung, P. (2006). Recombination mediator and Rad51 targeting activities of a human BRCA2 polypeptide. J Biol Chem 281, 11649-11657.

San Filippo, J., Sung, P., and Klein, H. (2008). Mechanism of eukaryotic homologous recombination. Annu Rev Biochem 77, 229-257.

Sarbajna, S., Davies, D., and West, S.C. (2014). Roles of SLX1-SLX4, MUS81-EME1, and GEN1 in avoiding genome instability and mitotic catastrophe. Genes Dev 28, 1124-1136.

Schwacha, A., and Kleckner, N. (1995). Identification of double Holliday junctions as intermediates in meiotic recombination. Cell 83, 783-791.

Seki, M., Nakagawa, T., Seki, T., Kato, G., Tada, S., Takahashi, Y., Yoshimura, A., Kobayashi, T., Aoki, A., Otsuki, M., et al. (2006). Bloom helicase and DNA topoisomerase IIIalpha are involved in the dissolution of sister chromatids. Mol Cell Biol 26, 6299-6307.

Shah Punatar, R., Martin, M.J., Wyatt, H.D., Chan, Y.W., and West, S.C. (2017). Resolution of single and double Holliday junction recombination intermediates by GEN1. Proc Natl Acad Sci U S A 114, 443-450.

Shinohara, A., Ogawa, H., and Ogawa, T. (1992). Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. Cell 69, 457-470.

Singh, T.R., Ali, A.M., Busygina, V., Raynard, S., Fan, Q., Du, C.H., Andreassen, P.R., Sung, P., and Meetei, A.R. (2008). BLAP18/RMI2, a novel OB-fold-containing protein, is an essential component of the Bloom helicase-double Holliday junction dissolvasome. Genes Dev 22, 2856-2868.

Sung, P. (1997). Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. J Biol Chem 272, 28194-28197.

Sung, P., and Robberson, D.L. (1995). DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. Cell 82, 453-461.

Svendsen, J.M., Smogorzewska, A., Sowa, M.E., O'Connell, B.C., Gygi, S.P., Elledge, S.J., and Harper, J.W. (2009). Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. Cell 138, 63-77.

Symington, L.S., and Gautier, J. (2011). Double-strand break end resection and repair pathway choice. Annu Rev Genet 45, 247-271.

Szakal, B., and Branzei, D. (2013). Premature Cdk1/Cdc5/Mus81 pathway activation induces aberrant replication and deleterious crossover. EMBO J 32, 1155-1167.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., and Stahl, F.W. (1983). The double-strand-break repair model for recombination. Cell 33, 25-35.

Tay, Y.D., and Wu, L. (2010). Overlapping roles for Yen1 and Mus81 in cellular Holliday junction processing. J Biol Chem 285, 11427-11432.

Tomlinson, C.G., Atack, J.M., Chapados, B., Tainer, J.A., and Grasby, J.A. (2010). Substrate recognition and catalysis by flap endonucleases and related enzymes. Biochem Soc Trans 38, 433-437.

Tsutakawa, S.E., Classen, S., Chapados, B.R., Arvai, A.S., Finger, L.D., Guenther, G., Tomlinson, C.G., Thompson, P., Sarker, A.H., Shen, B., et al. (2011). Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. Cell 145, 198-211.

Tsutakawa, S.E., Lafrance-Vanasse, J., and Tainer, J.A. (2014). The cutting edges in DNA repair, licensing, and fidelity: DNA and RNA repair nucleases sculpt DNA to measure twice, cut once. DNA Repair (Amst) 19, 95-107.

Wallis, J.W., Chrebet, G., Brodsky, G., Rolfe, M., and Rothstein, R. (1989). A hyper-recombination mutation in S. cerevisiae identifies a novel eukaryotic topoisomerase. Cell 58, 409-419.

Wang, C., Higgins, J.D., He, Y., Lu, P., Zhang, D., and Liang, W. (2017). Resolvase OsGEN1 Mediates DNA Repair by Homologous Recombination. Plant Physiol 173, 1316-1329.

Wechsler, T., Newman, S., and West, S.C. (2011). Aberrant chromosome morphology in human cells defective for Holliday junction resolution. Nature 471, 642-646.

West, S.C. (2009). The search for a human Holliday junction resolvase. Biochem Soc Trans 37, 519-526.

Whitby, M.C., Osman, F., and Dixon, J. (2003). Cleavage of model replication forks by fission yeast Mus81-Eme1 and budding yeast Mus81-Mms4. J Biol Chem 278, 6928-6935.

Wild, P., and Matos, J. (2016). Cell cycle control of DNA joint molecule resolution. Curr Opin Cell Biol 40, 74-80.

Wu, L., and Hickson, I.D. (2003). The Bloom's syndrome helicase suppresses crossing over during homologous recombination. Nature 426, 870-874.

Wyatt, H.D., Laister, R.C., Martin, S.R., Arrowsmith, C.H., and West, S.C. (2017). The SMX DNA Repair Tri-nuclease. Mol Cell 65, 848-860 e811.

Wyatt, H.D., Sarbajna, S., Matos, J., and West, S.C. (2013). Coordinated actions of SLX1-SLX4 and MUS81-EME1 for Holliday junction resolution in human cells. Mol Cell 52, 234-247.

Wyatt, H.D., and West, S.C. (2014). Holliday junction resolvases. Cold Spring Harb Perspect Biol 6, a023192.

Xu, D., Guo, R., Sobeck, A., Bachrati, C.Z., Yang, J., Enomoto, T., Brown, G.W., Hoatlin, M.E., Hickson, I.D., and Wang, W. (2008). RMI, a new OB-fold complex essential for Bloom syndrome protein to maintain genome stability. Genes Dev 22, 2843-2855.

Yamagata, K., Kato, J., Shimamoto, A., Goto, M., Furuichi, Y., and Ikeda, H. (1998). Bloom's and Werner's syndrome genes suppress hyperrecombination in yeast sgs1 mutant: implication for genomic instability in human diseases. Proc Natl Acad Sci U S A 95, 8733-8738.

Yang, H., Jeffrey, P.D., Miller, J., Kinnucan, E., Sun, Y., Thoma, N.H., Zheng, N., Chen, P.L., Lee, W.H., and Pavletich, N.P. (2002). BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. Science 297, 1837-1848.

Yang, H., Li, Q., Fan, J., Holloman, W.K., and Pavletich, N.P. (2005). The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA-ssDNA junction. Nature 433, 653-657.

Zellweger, R., Dalcher, D., Mutreja, K., Berti, M., Schmid, J.A., Herrador, R., Vindigni, A., and Lopes, M. (2015). Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. J Cell Biol 208, 563-579.

Zhu, Z., Chung, W.H., Shim, E.Y., Lee, S.E., and Ira, G. (2008). Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. Cell 134, 981-994.

Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisor Christian Biertümpfel, who offers me such great opportunity to work on this exciting project. I appreciate his belief and the guidance. His open-minded and fully supportive style always encourages and motivates me throughout my PhD study.

Furthermore, I would like to thank my thesis advisor Elena Conti for establishing the excellent research infrastructure and atmosphere in the department. Extended thanks also go to the other members of my Thesis Advisory Committee, Wolfgang Zachariae, Gyula Timinszky, and Boris Pfander for providing directions and valuable suggestions to my project.

I want to thank all the people in the Biochemistry Core Facility and the Crystallization Facility at the Max Planck Institute of Biochemistry for their analyses, experimental setups and the excellent technical support. My deep gratitude goes also to Jérôme Basquin for the great help in X-ray data acquisition and crystallographic analysis, Rajan Prabu for software supports and Claire Basquin, Ralf Stehle, André Mourão for sharing their knowledge of SAXS. I would also like to thank Naoko Mizuno and Esben Lorentzen for their informative comments and suggestions in lab meetings and sharing lab equipment on the second floor of the Department of Structural Cell Biology. Special thanks go to Bianca Habermann, Boris Pfander and Lissa Princz for the fruitful collaboration.

The International Max Planck Research School for Molecular Life Sciences (IMPRS-LS) provided interesting and structured lectures, workshops and organized many events to bring my fellow students together.

I would also like to thank my present and former labmates, Carina, Giulia, Maren, Marcus, Iuliia, Marvin, and Beth. It was really a good time working with all of you. Alongside with the Biertümpfel group, Dirk, Steffi, Qianmin, Nirakar, Melanie, Krissy, Julia, and Sagar, I want to thank you all for making me feel we are doing research in a big team and sharing the fun together during my working and non-working hours, early and late stage of the study.

Students Leo, Mohammed, Fanny, Ola, Chris, Felix, and Jonas, I thank you for bringing the vibrant and lively atmosphere to the lab, I benefited and learned a lot from all of you.

Finally, I want to thank my parents and family, without their understanding and continuous support, this thesis would not have been possible.