# 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] and Christian Biertümpfel[1 *]

[1] Max Planck Institute of Biochemistry, Department of Structural Cell Biology, Molecular Mechanisms of DNA Repair, Am Klopferspitz 18, 82152 Martinsried, Germany

[2] Max Planck Institute of Biochemistry, Department of Molecular Cell Biology, DNA Replication and Genome Integrity, Am Klopferspitz 18, 82152 Martinsried, Germany

[3] Max Planck Institute of Biochemistry, Computational Biology, Am Klopferspitz 18, 82152 Martinsried, Germany

[*] Corresponding Author
email: biertuempfel@biochem.mpg.de
phone: +49 89 8578 3641
fax: +49 89 8578 3605

18 # Abstract

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

32 # Introduction

33 Homologous recombination (HR) is a fundamental pathway ensuring genome integrity and
34 genetic variability (Heyer, 2015). In mitotic cells, double-strand breaks (DSBs) can be
35 repaired by HR using the sister chromatid as a template to restore the information in the
36 complementary double strand. In meiosis, the repair of programmed DSBs by HR and the
37 formation of crossovers are crucial to provide physical linkages between homologs and to
38 segregate homologous chromosomes. Furthermore, HR during meiosis creates sequence

39 diversity in the offspring through the exchange between homologs (Petronczki et al., 2003;
40 Sarbajna and West, 2014).
41 HR proceeds by pathways that may lead to the formation of DNA four-way junctions or
42 Holliday junctions (HJs) that physically link two homologous DNA duplexes (Heyer, 2015;
43 Holliday, 1964; Schwacha and Kleckner, 1995; Szostak et al., 1983). Faithful removal of HJs
44 is critical to avoid chromosome aberrations (Wechsler et al., 2011) and cells have evolved
45 sophisticated measures to disentangle joint molecules. One basic mechanism is resolution
46 mediated by HJ resolvases that introduce precise symmetrical nicks into the DNA at the
47 branch point. Nicked DNA strands are then rejoined by endogenous ligases leading to fully
48 restored or recombined DNA strands. This mechanism is well studied for bacterial and
49 bacteriophage resolvases such as *Escherichia coli* RuvC, T7 endonuclease I, T4
50 endonuclease VII (Benson and West, 1994; Lilley and White, 2001). These resolvases
51 operate as dimers and show a large degree of conformational flexibility in substrate
52 recognition and in aligning both active sites for coordinated cleavage. Interestingly, T4
53 endonuclease VII and RuvC reach into and widen the DNA junction point whereas T7
54 endonuclease I binds DNA by embracing HJs at the branch point (Biertümpfel et al., 2007;
55 Górecka et al., 2013; Hadden et al., 2007).
56 In eukaryotes, HR is more complex and tightly regulated. In somatic cells, HJ
57 dissolution by a combined action of a helicase and a topoisomerase (BLM-TOPIIIα-RMI1-
58 RMI2 complex in humans) is generally the favored pathway, possibly to restore the original
59 (non-crossover) DNA arrangement (Cejka et al., 2010, 2012; Ira et al., 2003; Putnam et al.,
60 2009; Wu and Hickson, 2003). In contrast, HJ resolution generates crossover and non-
61 crossover arrangements depending on cleavage direction. Several endonucleases such as
62 GEN1, MUS81-EME1 and SLX1-SLX4 have been implicated as HJ resolvases in eukaryotes

63 (Andersen et al., 2011; Castor et al., 2013; Fekairi et al., 2009; Garner et al., 2013; Ip et al.,
64 2008; Muñoz et al., 2009; Svendsen and Harper, 2010; Svendsen et al., 2009; Wyatt et al.,
65 2013). Interestingly, these resolvases are not structurally related and have different
66 domain architectures, giving rise to variable DNA recognition and regulation mechanisms.
67 The interplay between resolution and dissolution mechanisms is not fully understood yet,
68 however, cell cycle regulation of resolvases seems to play an important role (Blanco et al.,
69 2014; Chan and West, 2014; Eissler et al., 2014; Matos et al., 2011).

70 GEN1 belongs to the Rad2/XPG family of structure-selective nucleases that are
71 conserved from yeast to humans (Ip et al., 2008; Lieber, 1997; Yang, 2011). The Rad2/XPG
72 family has four members with different substrate preferences that function in DNA
73 maintenance (Nishino et al., 2006; Tsutakawa et al., 2014). They share a conserved N-
74 terminal domain (XPG-N), an internal domain (XPG-I) and a 5'->3' exonuclease C-terminal
75 domain containing a conserved helix-hairpin-helix motif. C-terminal to the nuclease core is
76 a regulatory region that is diverse in sequence and predicted to be largely unstructured.
77 Although the catalytic cores are well conserved in the superfamily, substrate recognition is
78 highly diverse: XPG/Rad2/ERCC5 recognizes bubble/loop structures during nucleotide-
79 excision repair (NER), FEN1 cleaves flap substrates during Okazaki fragment processing in
80 DNA replication, EXO1 is a 5'->3' exonuclease that is involved in HR and DNA mismatch
81 repair (MMR) and GEN1 recognizes Holliday junctions (Grasby et al., 2012; Ip et al., 2008;
82 Nishino et al., 2006; Tomlinson et al., 2010; Tsutakawa et al., 2014). A common feature of
83 the superfamily is their inherent ability to recognize flexible or bendable regions in the
84 normally rather stiff DNA double helix. Interestingly, GEN1 shows versatile substrate
85 recognition accommodating 5' flaps, gaps, replication fork intermediates and Holliday
86 junctions (Ip et al., 2008; Ishikawa et al., 2004; Kanai et al., 2007). According to the current

87 model, however, the primary function of GEN1 is HJ resolution (Garner et al., 2013;
88 Sarbajna and West, 2014; West et al., 2015) and it is suggested to be a last resort for the
89 removal of joint molecules before cytokinesis (Matos et al., 2011).

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

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

110 **Results**

111 **Structure determination and architecture of the GEN1-DNA complex**

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

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

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

151 **GEN1 has a conserved chromodomain with a closed aromatic cage**

152 Chromodomains are found in many chromatin-associated proteins that bind modified
153 histone tails for chromatin targeting (reviewed in Blus et al., 2011; Eissenberg, 2012; Yap
154 and Zhou, 2011), but it has not previously been associated with nucleases. To understand
155 the significance of the chromodomain for the function of GEN1, we first examined if the
156 chromodomain is conserved in GEN1 homologs using HMM-HMM (Hidden Markov Models)
157 comparisons in HHPRED (Söding et al., 2005). We found that the chromodomain in GEN1 is

158 conserved from yeast (Yen1) to humans (Figure 3A). The only exception is *Caenorhabditis*
159 *elegans* GEN1, which has a much smaller protein size of 443 amino acids compared to yeast
160 Yen1 (759 aa) or human GEN1 (908 aa).

161 To further compare the structural arrangement of the aromatic cage in human GEN1
162 with other chromodomains we analyzed the best matches from the DALI search (Figure
163 3B). We found many hits for different chromo- and chromo-shadow domains with root
164 mean square deviations between 1.9 and 2.8 Å (compare Figure 3-figure supplement 1). A
165 superposition of the aromatic cage of the five structurally most similar proteins with GEN1
166 (Figure 3C) showed that residues W418, T438 and E440 are well conserved, whereas two
167 residues at the rim of the canonical binding cleft are changed from phenylalanine/tyrosine
168 to a leucine (L397) in one case and a proline (P421) in another (Figure 3C). Instead, Y424
169 occupies the space proximal to P421, which is about 1.5 Å outwards of the canonical cage
170 and widens the GEN1 cage slightly. The substitution of phenylalanine/tyrosine to leucine is
171 also found in CBX chromo-shadow domains (see below); however, the rest of the GEN1
172 aromatic cage resembles rather chromodomains.

173 Chromodomains often recognize modified lysines through their aromatic cage thus
174 targeting proteins to chromatin (reviewed in Blus et al., 2011; Eissenberg, 2012; Yap and
175 Zhou, 2011). Given the conserved aromatic cage in GEN1 we tested the binding to modified
176 histone tail peptides (Figure 3C/D). However, we did not detect any binding despite
177 extensive efforts using various histone tail peptides in pull-down assays, microscale-
178 thermophoresis (MST) or fluorescence anisotropy measurements (compare Figure 3-figure
179 supplement 3 and 4). Our structure shows that the aromatic cage is closed by helix α15
180 (Figure 3E blue/pink), which has a hydrophobic interface towards the aromatic cage with
181 residues L376, T380 and M384 reaching into it (compare Figure 4F). This potentially

182 hampers the binding of the tested peptides in this conformation under physiological
183 conditions.

184 **The GEN1 chromodomain is distantly related to CBX and CDY chromodomains**

185 To explore the functional role of the GEN1 chromodomain, we evaluated its similarity to
186 other chromodomains by comparing all of the 46 known human chromodomains from 34
187 different proteins. We made pairwise comparisons with HHPRED, PSIBLAST, combined the
188 alignments and generated a phylogenetic tree (Figure 3F and Figure 3-figure supplement
189 2). The analysis showed a tree branching into known subfamilies: chromobox proteins
190 (CBX, red), chromodomain Y-linked proteins (CDY, yellow), chromodomain-helicase DNA-
191 binding proteins (blue) and chromo-barrel domain proteins (green). The GEN1
192 chromodomain was found to be distantly related to the CDY chromodomains and
193 chromobox proteins, particularly to the chromo-shadow domains of CBX1, CBX3 and CBX5.
194 This agrees with the result from the DALI search, in which CBX chromo-shadow domains
195 and homologs thereof were among the closest structural matches. Together with the
196 observed differences in residues forming the aromatic cage, it indicates that the GEN1
197 chromodomain forms a new subgroup with features from chromo- and chromo-shadow
198 domains that emerged from a common ancestor within CBX/CDY proteins.

199 **GEN1-DNA interactions**

200 The GEN1-HJ structure revealed that the upstream DNA-binding interface acts as a docking
201 site for double-stranded DNA and that the chromodomain secures its position. The DNA is
202 bound at the upstream interface and the hydrophobic wedge but does not extend into the
203 active site or to the downstream interface (Figure 1B/C/D). Comparison of the structure of

204 GEN1 to related structures of FEN1, Rad2 and EXO1 (Miętus et al., 2014; Orans et al., 2011;
205 Tsutakawa et al., 2011) suggests that a DNA substrate has to extend to the downstream
206 interface to position a DNA strand for cleavage by the active site of GEN1 (Figure 1B/C and
207 Figure 1F). In the GEN1 structure, the end of the DNA arm attaches to the hydrophobic
208 wedge provided by parts of helices α2-α3 and their connecting loop (Figure 4A/B), forming
209 van-der-Waals contacts with the first base pair, which docks perfectly onto the protruding
210 curb of residues 41-51 (Figure 4B). The uncleaved DNA strand is further stabilized and its
211 geometrical arrangement is fixed by the upstream DNA-binding interface. Particularly, the
212 DNA is contacted by a β-pin (strands β6-β7; Figure 4A/C) from one side and by R54 and
213 F58 (Figure 4A/D) from helix α3 together with Y370 and K374 (helix α15) from the
214 opposite side (Figure 4A/C). The key residues in the β-pin are T171 that forms a hydrogen
215 bridge to the phosphate of the first base (Figure 4A, "G1") and M172 that makes a van-der-
216 Waals contact to the DNA backbone at the second base (Figure 4A, "A2"). R54 reaches into
217 the DNA minor groove and forms a hydrogen bond with the ribose ring oxygen at the third
218 base of the uncleaved strand and F58 packs against the same ribose moiety (Figure 4C/D).
219 Y370 and K374 in α15 form hydrogen bonds to the backbone of the third base of the
220 uncleaved DNA strand (Figure 4D, "G3").

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

228 and the nuclease core (Figure 4E). The interactions at the chromodomain extend the
229 upstream DNA-binding interface to cover a full DNA turn, reinforcing the binding.

230 The downstream binding interface can be inferred from other Rad2/XPG structures
231 (Figure 1C/F) as the nuclease core is well conserved in GEN1, FEN1, Rad2 and EXO1 (root
232 mean square deviations of 0.9-1.1 Å for 161 Cα atoms, respectively). The residues
233 corresponding to the tip of the thumb (residues 79-92), which are disordered in the GEN1
234 structure, likely form helix α4 upon DNA binding to the downstream interface as seen in
235 human FEN1 and EXO1 (Orans et al., 2011; Tsutakawa et al., 2011). The missing residues in
236 GEN1 have 35.7% identity and 78.6 % similarity (BLOSUM62 matrix) to the corresponding
237 residues in FEN1 (90-103), which form helix α4 in the FEN1-DNA complex (compare
238 Figure 2). The same region is disordered in FEN1 when no DNA is bound (Sakurai et al.,
239 2005). This indicates that also GEN1 undergoes such a disorder-to-order transition to form
240 an arch with helices α4 and α6 upon substrate binding (Patel et al., 2012) and similar to the
241 arrangement in T5 FEN (Ceska et al., 1996).

242 **The activity of GEN1 depends on correct DNA positioning.**

243 GEN1 has versatile substrate recognition features, ranging from gaps, flaps, replication fork
244 intermediates to HJs (Ip et al., 2008; Ishikawa et al., 2004; Kanai et al., 2007). To
245 understand the functional relevance of the GEN1 structure for DNA recognition we
246 performed a series of mutagenesis studies with single point mutations and truncated
247 protein variants (Figure 5 and Figure 5-figure supplement 2/3) to investigate the effect on
248 the active site (D30N), upstream DNA binding (R54E), downstream DNA binding (C36E),
249 arch at the downstream interface (R89E, R93E, H109E, F110E), and chromodomain
250 (Δchromo, K404E, R406E). We performed nuclease assays by titrating different amounts of

251 GEN1 to a fixed DNA concentration of 40 nM for 15 min and DNA cleavage products were
252 analyzed by native electrophoresis (Figure 5A and Figure 5-figure supplement 2/3). We
253 used an immobile HJ and a 5' flap substrate side-by-side to facilitate the comparison of the
254 effects on separate GEN1 functions. Notably, stoichiometric amounts of GEN1 were
255 required to cleave HJ substrates whereas 5' flaps were readily processed with catalytic
256 amounts (Figure 5A).

257 The active site modification D30N showed that the cleavage activity on both HJ and 5'
258 flap substrates was lost in agreement with previously published data (Ip et al., 2008).
259 According to our structure, R54 in helix α3 at the upstream interface fixes the substrate
260 position by reaching into the minor DNA groove and we observed that R54E had a strongly
261 reduced cleavage activity (~50%; Figure 5B), indicating a key role in substrate positioning.

262 Residue C36 in helix α2 points towards the downstream interface and likely contacts
263 the DNA upon binding (compare Figure 5D). The corresponding FEN1 Y40, is a key residue
264 stacking with the -1 base of the 5' flap at the FEN1 active site (Tsutakawa et al., 2011).
265 Therefore, we tested the cleavage ability of a $GEN1^{C36E}$ and found that the mutant protein
266 had completely lost its enzymatic activity for both, HJ and 5' flap cleavage, to the same
267 degree as the active site modification D30N (Figure 5B). This effect is stronger than for
268 $FEN1^{Y40A}$, which showed only a partial loss in activity (Tsutakawa et al., 2011). Our results
269 suggest that C36 provides a polar interface for orienting and guiding the cleaved strand
270 towards the active site and the lower gateway.

271 We further tested a glutamate modification of the superfamily-conserved R89 and R93
272 located in the disordered part continuing to helix α6, presumably forming an arch (see
273 above). The arch was shown to facilitate cleavage by clamping flap substrates in FEN1 and
274 the modification R100A showed a strong decrease in the cleavage activity (Patel et al.,

275 2012). The GEN1 R89E mutation, corresponding to residue R100 in FEN1, showed that the
276 activity of GEN1 with a HJ substrate was not altered. In the case of a 5' flap substrate,
277 cleavage was slightly reduced and it reached to the full level at enzyme concentrations
278 higher than 10 nM. The effect of the R93E modification was even less pronounced
279 compared to R89E. In contrast, the cleavage of both 5' flap and HJ substrates depended
280 strongly on F110 at helix α6 (thumb), which points towards the active site. An F110E
281 modification showed a reduction in cleavage by 25% for HJ substrates and the effect was
282 even stronger for 5' flap substrates, where the activity is reduced by 65%. The equivalent
283 position in FEN1 is V133 showing a critical involvement in stabilizing 5' flap DNA by
284 orienting the -1 nucleotide for catalysis (Tsutakawa et al., 2011). We have also tested the
285 effect of modifying H109, which neighbors the critical F110. Even though it points away
286 from the active site, a glutamate at this position reduced 5' flap cleavage to 83% and HJ
287 cleavage recovered only at high substrate concentrations of 256 nM. Overall, the results
288 suggest that F110 has a key position for DNA recognition and processing.

289 **Coordinated Cleavage of HJs**

290 Classical HJ resolvases introduce two symmetrical incisions across the junction point by
291 coordinating the action of two active sites. The first nick is rate-limiting and the second one
292 takes place near-simultaneously and within the lifetime of the resolvase-DNA complex.
293 This mechanism has been well studied for bacterial and bacteriophage HJ resolvases (Fogg
294 and Lilley, 2000; Giraud-Panis and Lilley, 1997; Pottmeyer and Kemper, 1992; Shah et al.,
295 1997). Hence, it is thought that also GEN1 dimerizes upon binding to HJ substrates as
296 indicated by coordinated cleavage and by an increase in hydrodynamic radius compared to
297 protein alone (Chan and West, 2015; Rass et al., 2010). In order to further examine the

298 effect of GEN1 modifications on HJ cleavage we used a cruciform plasmid cleavage assay to
299 evaluate GEN1's nicking function, as illustrated in Figure 5E. Here, the plasmid pIRbke8$^{mut}$
300 served as a substrate that contains an inverted-repeat sequence extruding a cruciform
301 structure when supercoiled (Chan and West, 2015; Lilley, 1985; Rass et al., 2010).
302 Coordinated dual incision of the cruciform, i.e. by a dimer, leads to linear duplex products
303 with slow migration, whereas uncoordinated cleavage, i.e. monomeric enzymes, results in
304 nicked plasmids that migrate even slower (Figure 5F). Cruciform structures are reabsorbed
305 when the superhelical stress is released upon single nicking and the DNA cannot serve as a
306 substrate anymore.

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

318 **The Chromodomain of GEN1 Facilitates Efficient Substrate Cleavage**

319 Agreeing with the structural significance for DNA binding, the truncation of the
320 chromodomain (Δchromo, residues 2-389) showed a severe reduction (~3-fold) in HJ

321 cleavage activity whereas all longer GEN1 fragments containing the chromodomain (2-464,
322 2-505 and 2-551) showed full activity (Figure 5-figure supplement 3). Interestingly, the
323 effect of the chromodomain truncation is even more pronounced for 5' flap DNA cleavage
324 than for HJs, showing a 7-fold reduction compared to wild type (Figure 5C). The activity of
325 GEN1 in the plasmid-based cruciform cleavage assay was also severely hampered in the
326 absence of the chromodomain (Figure 5F) showing only a weak band for linear products
327 and no increase for nicked plasmid, emphasizing the importance of the chromodomain for
328 GEN1 activity.

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

337 **Influence of Phosphorylation-Mimicking Chromodomain Modifications**

338 PhosphoSitePlus (Hornbeck et al., 2014) lists two phosphorylation sites at residues T380
339 and T438 in GEN1 that were found in a T-cell leukemia and a glioblastoma cell line. These
340 residues are located in helix α15 and at the rim of the aromatic cage, respectively. Both
341 phosphorylation sites are positioned to interrupt hydrophobic interactions between helix
342 α15 and the chromodomain (Figure 5D and Figure 4F). Therefore, we tested if the
343 phosphorylation-mimicking modifications T380E and T438E had an effect on GEN1's

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

350 **Physiological Relevance of GEN1 interactions**

351 To test the physiological relevance of the identified GEN1-DNA interactions, we
352 investigated the survival of *Saccharomyces cerevisiae* mutant strains expressing variants of
353 Yen1 (GEN1 homolog) after treatment with the DNA-damaging agent MMS (Figure 5G and
354 Figure 5-figure supplement 4/5). All Yen1 variants were expressed to a similar degree as
355 endogenous Yen1, which was confirmed by Western Blot analysis (Figure 5-figure
356 supplement 4). Because of the functional overlap of Mus81 and Yen1 in HR (Blanco et al.,
357 2010) a double knockout (*yen1Δ mus81Δ*) was used and complemented with different
358 variants of *Yen1*.

359 The control strain, complemented with wild type Yen1, survived MMS concentrations
360 of up to 0.01%, consistent with the described hypersensitivity of *mus81Δ* mutants (Blanco
361 et al., 2010; Interthal and Heyer, 2000). In stark contrast, cells containing either the active
362 site mutant Yen1-D41N (corresponding to $GEN1^{D30N}$) or the downstream interface mutant
363 Yen1-F47E (corresponding to $GEN1^{C36E}$) did not grow even at an MMS concentration as low
364 as 0.0025% (Figure 5G). After expression of the upstream interface mutant Yen1-I97E
365 (corresponding to $GEN1^{R54E}$) cells showed a slight but significant growth defect at high
366 MMS concentrations (see panels for 0.0075% and 0.01% MMS in Figure 5G). These results

367 are therefore consistent with the *in vitro* cleavage results carried out with GEN1 mutants
368 and showing a reduction in activity for R54E and no activity for C36E (see Figure 5C). As a
369 last mutant in the nuclease core, we tested the K298E mutation which is located in helix
370 α10 of the H2TH motif in the downstream DNA-binding interface, and for which we were
371 unable to obtain the corresponding GEN1$^{K219E}$ modification for cleavage assays (compare
372 Figure 5D). This mutant displayed a strong sensitivity towards MMS but lower than the one
373 observed for the catalytic mutant, indicating that the mutant was partially functional in
374 yeast (Figure 5G).

375 We next investigated the effect of mutations in the aromatic cage of Yen1's
376 chromodomain (compare Figure 3) and found that their severity was strongly position
377 dependent. Mutation of R486E and Y487A in Yen1, both of which are located near the base
378 of the cage, corresponding to the W418 position in GEN1 (see Figure 3C), showed a strong
379 effect on MMS sensitivity (see Figure 5G), similar to the one observed for the catalytic
380 mutant, presumably due to a dysfunctional chromodomain. In contrast, mutations located
381 further outside of the core (F478A and K484E*)* led to a less pronounced MMS sensitivity.
382 The same was true for the K469E variant, which corresponds to position R406 at the
383 chromodomain-DNA interface in GEN1 (see Figure 3A and 5F), and for residues at the rim
384 of the chromodomain (*yen1-N526A, yen1-L528D* and *yen1-W529A*), consistent with our *in*
385 *vitro* observation for GEN1$^{T438E}$ (slightly reduced activity, Figure 5C). No effect on MMS
386 sensitivity was detected for *yen1-L530A*, which corresponds to a conserved glutamate in
387 chromodomains (E440 in GEN1). Lastly, we found that the deletion of the chromodomain
388 (Yen1-Δ452-560) lead to a severe phenotype comparable to the active site mutant Yen1-
389 D41N (Figure 5G and Figure 5 complement 5). The Yen1 variant lacking the chromodomain
390 was expressed to levels similar to the full-length protein and we therefore conclude that

391 the chromodomain is crucial for the function of Yen1. Taken together the functional data of
392 Yen1 mutants *in vivo* and GEN1 mutants *in vitro* point towards an essential and
393 evolutionary conserved role of the chromodomain in GEN1/Yen1 proteins.

# 394 Discussion

## 395 Implications of the Chromodomain

396 The structure of the human GEN1 catalytic core provides the missing structural
397 information in the Rad2/XPG family. The GEN1 structure complements recent reports on
398 the structures of Rad2, EXO1 and FEN1, (Miętus et al., 2014; Orans et al., 2011; Tsutakawa
399 et al., 2011). Thereby, it gives insights how relatively conserved nuclease domains
400 recognize diverse substrates in a structure-selective manner and act in different DNA
401 maintenance pathways. In comparison with other Rad2/XPG nucleases, GEN1 shows many
402 modifications on common structural themes that give the ability to recognize a diverse set
403 of substrates including replication fork intermediates and HJs. The upstream DNA interface
404 of GEN1 lacks the "acid block" found in FEN1, instead it has a prominent groove at the same
405 position (compare Figure 1, "upper gate") with a strategically positioned R54 nearby.
406 Furthermore, the helical arch in GEN1 misses helix α5, which forms a cap structure in FEN1
407 and EXO1 that stabilizes 5' overhangs for cleavage. These features have implications for the
408 recognition and cleavage of HJ substrates (see below). The most striking difference to other
409 Rad2/XPG family members is that the GEN1 nuclease core is extended by a chromodomain,
410 which provides an additional DNA anchoring point for the upstream DNA-binding
411 interface. The evolutionarily conserved chromodomain is important for efficient substrate
412 cleavage as we showed using truncation and mutation analyses. This finding opens new

413 perspectives for the regulation of GEN1 and for its interactions with other proteins.
414 Chromodomains serve as chromatin-targeting modules (reviewed in Blus et al., 2011;
415 Eissenberg, 2012; Yap and Zhou, 2011), general protein interaction elements (Smothers
416 and Henikoff, 2000) as well as dimerization sites (Canzio et al., 2011; Cowieson et al., 2000;
417 Li et al., 2011). These possibilities are particularly interesting, as chromatin targeting of
418 proteins via chromodomains has been implicated in the DNA damage response. The
419 chromatin remodeler CHD4 is recruited in response to DNA damage to decondense
420 chromatin (reviewed in O'Shaughnessy and Hendrich, 2013; Stanley et al., 2013). The
421 chromodomains in CHD4 distinguish the histone modifications H3K9me3 and H3K9ac and
422 determine the way how downstream DSB repair takes place (Ayrapetov et al., 2014; Price
423 and D'Andrea, 2013). It is plausible that GEN1 uses its chromodomain not only as a
424 structural module to securely bind DNA but also for targeting or regulatory purposes. Even
425 though it was not possible to find any binding partner with a series of tested histone tail
426 peptides, we cannot exclude that the chromodomain is used as an interaction motif or
427 chromatin reader. It will therefore be interesting to extend our interaction analysis to a
428 larger number of peptides and proteins. Interestingly, the modifications GEN1$^{L397E}$ and
429 GEN1$^{Y424A}$ at the rim of the chromodomain did not alter DNA cleavage activity (Figure 5-
430 figure supplement 2), however, mutations of residues at the rim of Yen1's chromodomain
431 show a phenotype, suggesting an additional role like binding to an endogenous factor.

432 Another intriguing aspect of the chromodomain is that the conserved T438 at the rim
433 of the aromatic cage and T380 at the closing helix α15 are both part of a casein kinase II
434 consensus sequence for phosphorylation (Ser/Thr-X-X-Asp/Glu). Ayoub et al., 2008
435 showed that the analogous threonine in the chromodomain of CBX1 is phosphorylated in
436 response to DNA damage and phosphorylation disrupts the binding to H3K9me. We

437 observed a reduction in DNA cleavage activity for the phosphorylation mimicking
438 mutations T380E and T438E, which may suggest a regulatory role. They might function
439 together and in combination with other modifications to provide a way of functional
440 switching at the chromodomain. Furthermore, Blanco et al., 2014 and Eissler et al., 2014
441 recently identified several CDK phosphorylation sites in an insertion in the Yen1
442 chromodomain which affects HJ cleavage and together with phosphorylation of a nuclear
443 localization signal (NLS) in the regulatory domain restricts Yen1's activity to anaphase. The
444 insertion is not found in other chromodomains and it is extended in Yen1 compared to
445 GEN1, which is lacking these phosphorylation sites (compare Figure 3A/B). Notably, the
446 activity of Yen1 is negatively regulated by CDK-dependent phosphorylation (Blanco et al.,
447 2014; Chan and West, 2014; Eissler et al., 2014; Matos et al., 2011), suggesting that the
448 chromodomain is targeted by cell cycle kinases. It also provides a likely explanation for the
449 different regulatory mechanisms found in GEN1 and Yen1 (Blanco and Matos, 2015; Chan
450 and West, 2014; Matos and West, 2014). Exploration of the regulatory function of the GEN1
451 chromodomain will be an important topic to follow up and this may lead to the
452 understanding of the precise regulation mechanism of GEN1 as well as its substrate
453 recognition under physiological conditions.

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

457 **Recognition of DNA substrates**

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

460 substrate recognition of GEN1 (Figure 6). This was done by superimposing the protein part
461 of the FEN1-DNA complex (PDB 3q8k) onto our GEN1 structure and extending the DNA
462 accordingly (Figure 6A/B). Remarkably, the superimposition of the proteins aligns the DNA
463 from the FEN1 structure in the same register as the DNA in the GEN1 complex at the
464 upstream interface (Figure 6A and 6B insert). Furthermore, the free 5' and 3' ends of the
465 double flap DNA from the FEN1 structure point towards the lower and the upper gateway
466 in GEN1, respectively (Figure 6B). We extended the GEN1 structure by homology modeling
467 of the disordered residues 79-92 (helix α4) in GEN1 (Figure 6B). In addition to the
468 similarity of this part to FEN1, the model readily showed the arrangement forming an arch
469 structure. This would explain why GEN1 recognizes 5' flap substrates efficiently, analogous
470 to FEN1, as the arch can clamp a single-stranded DNA overhang for productive cleavage.
471 This also explains why the F110E modification in the arch at helix α6 hampered 5' flap
472 cleavage severely. The side chain points directly towards the active site and likely disturbs
473 the stabilization of a 5' overhang for catalysis by charge repulsion. However, there are two
474 features in GEN1 that vary from the arrangement in FEN1 and EXO1 considerably. Helix α6
475 is longer (24 instead of 15 residues) and helix α5 is missing in GEN1. As a result the arch
476 points away from the DNA rather than forming a "cap" structure as it is observed in FEN1
477 and EXO1 (Orans et al., 2011; Tsutakawa et al., 2011). Furthermore, the modified arch in
478 GEN1 provides an opening, marked as "lower gate" in Figure 6B. These differences are
479 likely the basis for GEN1's versatile DNA recognition features.

480 **Implications of an adjustable hatch in GEN1 for substrate discrimination**

481 The diverging orientation of the arch (helices α4 and α6) in GEN1 compared to the one in
482 FEN1 and EXO1 (helices α4, α5 and α6) may have thus significance for the recognition of HJ

483 substrates. By pointing away from the active site the arch provides an opening to
484 accommodate unpaired, single-stranded DNA to pass along the arch at the lower gate
485 (groove between α2 and α4) (Figure 6B "lower gate") from one GEN1 monomer to the
486 upper gate (groove between α2-α3 and α14) (Figure 6B "upper gate") of the other within a
487 GEN1 dimer (Figure 6B/C). R54 is perfectly positioned at the minor groove to guide the
488 second cleavage strand to pass through the upper gate (compare Figure 4 and Figure 6B/C,
489 marked with a asterisk). In FEN1, this position is occupied by the "acid block", which
490 stabilizes a single 3' flap of the unpaired substrate (Tsutakawa et al., 2011) and it would
491 not accommodate longer 3' DNA overhangs. In our model two GEN1 monomers come
492 together crosswise upon HJ binding (Figure 6C). The helical arches of both proteins likely
493 provide additional protein-protein interactions as well as protein-DNA contacts by packing
494 against the backbone of opposite DNA arms (Figure 6C). As a result, the GEN1 dimer
495 orients both active sites symmetrically across the junction point resembling the situation in
496 bacterial RuvC (Figure 6D; Bennett and West, 1995a; Górecka et al., 2013). This
497 arrangement would ensure that both incisions are introduced within the lifetime of the
498 GEN1-HJ complex as observed biochemically by us and others (Rass et al., 2010). The
499 mechanism likely works in a coordinated nick-and-counter-nick fashion, as shown for
500 bacterial or bacteriophage HJ resolvases (Fogg and Lilley, 2000; Giraud-Panis and Lilley,
501 1997; Pottmeyer and Kemper, 1992; Shah et al., 1997) and recently for GEN1 (Chan and
502 West, 2015).

503 The distance between both gates is bridged by unpaired bases in our GEN1-HJ model.
504 This view is supported by the observation that FEN1 unpairs two bases near the active site
505 through interactions with the hydrophobic wedge leading to strongly bent DNA arms
506 between the upstream and downstream DNA interfaces. This mechanism seems to be a

507 common feature of Rad2/XPG nucleases (Finger et al., 2013; Grasby et al., 2012; Tsutakawa
508 et al., 2011). Consistent with this view, the bacterial RuvC resolvase (Figure 6D) has also
509 been shown to unfold HJ junctions (Bennett and West, 1995b; Górecka et al., 2013). In the
510 case of GEN1, the critical step would be the assembly of the dimer around the junction
511 point in a highly restraint way and the introduction of the first nick. This releases the
512 tension on the complex like a spring leading to an immediate second cut and subsequent
513 disassembly of the GEN1-HJ complex. Furthermore, a HJ does not provide free DNA ends
514 and adopts a structure that intrinsically restrains its degrees of freedom, thus inhibiting
515 cleavage by a single GEN1 monomer. Altogether we speculate that the arch (helix α4-α6)
516 acts like a lever or hatch switching between flap and HJ recognition modes. When a free 5'
517 end is available it closes and clamps the flap, thus positions the DNA for cleavage. For the
518 case of a HJ substrate, the arch adopts an open conformation, allowing unpaired, single-
519 stranded DNA to pass, while preventing the correct positioning of the DNA for catalysis at
520 first. HJ cleavage is inhibited until a second GEN1 monomer binds. This mechanism differs
521 from the one used by bacterial or bacteriophage HJ resolvases, which act as obligate dimers
522 binding to DNA substrates in a concerted way (compare Figure 6D-F). Our model for DNA
523 cleavage by GEN1 describes a conformational switch provided by a flexible arch that can
524 discriminate between substrates containing free 5' ends or those with a restraint structure
525 like HJs. This aspect may explain our observation that GEN1 cleaves 5' flap DNA
526 catalytically while stoichiometric amounts are required for HJ substrates (Figure 5A-C).
527 Using a switchable hatch in a spring-loaded mechanism would be an efficient way of
528 preventing a single cut at a HJ junction while allowing GEN1 to adapt to recognize various
529 DNA substrates and perform different functional roles. Thus, GEN1 may have an intrinsic

530 safety mechanism that ensures symmetrical dual incision across a branch point. Further
531 studies have to address the exact engagement mechanism.

532 **GEN1 in a biological context**

533 GEN1's biological role is not fully understood yet. Yeast cells are viable without the GEN1
534 homolog Yen1 even in the presence of DNA damaging agents as the Mus81-Eme1 complex
535 can complement the defect (compare Figure 5-figure supplement 4; Blanco et al., 2010).
536 Consistently, both proteins can cleave 5' flaps and HJ substrates *in vitro*. However, GEN1
537 can cleave intact HJs symmetrically whereas MUS81-EME1 is much more efficient with
538 nicked DNA four-way junctions (Castor et al., 2013; Wyatt et al., 2013). Matos et al., 2011
539 suggested that Yen1/GEN1 might serve as a backup enzyme to resolve persistent HJs that
540 have eluded other mechanisms of joint molecule removal before cytokinesis.

541 Our analysis infers that HJ cleavage is slower than 5' flap cleavage (Figure 5B/C),
542 bringing interesting implications for a safety control of GEN1's activity. GEN1 may have to
543 assemble in an accurate way before it can cleave a HJ. Likewise, it increases GEN1's
544 persistence time on HJs and opens a window for branch migration for extending the length
545 of recombined stretches of DNA. Moreover, GEN1 recognizes various DNA substrates,
546 which may point towards a general role in processing substrates in different DNA
547 maintenance pathways. GEN1 has been shown to cleave replication fork intermediates and
548 it is implicated in the resolution of replication-induced HJs (Garner et al., 2013; Sarbajna et
549 al., 2014). Like MUS81-EME1, it might also be important for the processing of fragile sites
550 to ensure proper chromosome segregation (Ying et al., 2013). These functions have to be
551 tested systematically to understand GEN1's biological role. In this context, the regulation of
552 GEN1 is an important factor and needs to be explored. Our study identified a

553 chromodomain extending the GEN1 nuclease core that might have a role in regulating the
554 enzyme. An open question is the function and architecture of the remaining 444 amino
555 acids at the C-terminus of GEN1. They are thought to regulate the nuclease activity and
556 control subcellular localization (Blanco et al., 2014; Chan and West, 2014; García-Luis et al.,
557 2014). It is very likely that new interaction sites and post-translational modifications in
558 this region will be discovered in future. The presented structure together with additional
559 studies will help to unravel these questions and to obtain a comprehensive view of the
560 functions of the Rad2/XPG nucleases.

# 561 Materials and methods

# 562 Experimental Procedures

### 563 Protein expression and purification

564 Wild type human GEN1 and truncations thereof (residues 2-551, 2-505, 2-464, 2-389) were
565 amplified by PCR from IMAGE clone 40125755 (Mammalian Gene collection, natural
566 variant S92T, S310N, UniProtID Q17RS7) and cloned into a self-made ligation-independent
567 cloning vector with various C-terminal tags followed by -His8. Truncated versions were
568 designed based on limited proteolysis in combination with domain prediction and
569 functional assays to determine the smallest yet active fragment. The N-terminal methionine
570 was cleaved by cellular methionyl-aminopeptidase, which is an essential requirement in
571 the Rad2/XPG family as the N-terminus (conserved residue G2) folds towards the active
572 site. Mutations were introduced by site-directed mutagenesis using Phusion Polymerase
573 (NEB, Frankfurt/Main, Germany). All recombinant proteins were expressed in the *E. coli*
574 BL21(DE3) pRIL strain (MerckMillipore, Darmstadt, Germany). Cells were grown at 37°C

575 until mid-log phase and induced overnight with 0.2 mM IPTG at 16°C. Cells were harvested
576 by centrifugation and resuspended in lysis buffer containing 1x phosphate buffered saline
577 (PBS) with additional 500 mM NaCl, 10% (v/v) glycerol, 2 mM DTT, 1 mM EDTA, 1 μM
578 leupeptin, 1 μM pepstatin A, 0.1 mM AEBSF and 2 μM aprotinin and lyzed by sonication.
579 Cell debris was removed by centrifugation (75 600 *g* for 45 min), the clarified lysate was
580 applied onto Complete HisTag Nickel resin (Roche Diagnostics, Mannheim, Germany) and
581 washed with buffer A consisting of 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 10% (v/v)
582 glycerol, 2 mM DTT and followed by a chaperone wash step with 20 mM Tris-HCl pH 7.5,
583 500 mM NaCl, 2 mM ATP, 5 mM $MgCl_2$, 10% (v/v) glycerol and 2 mM DTT. The protein was
584 eluted with buffer A containing 300 mM imidazole. The tag was cleaved, followed by cation
585 exchange chromatography using a HiTrap SP HP column (GE Healthcare, Freiburg,
586 Germany) with a linear gradient from 150 mM to 450 mM NaCl. Peak fractions were pooled
587 and further purified by size-exclusion chromatography on a HiLoad 16/60 Superdex 200
588 (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 100 mM NaCl, 5%(v/v) glycerol,
589 0.1 mM EDTA and 2 mM TCEP. Peak fractions were pooled, concentrated, flash-frozen in
590 liquid nitrogen and stored at -80°C.

591 **Crystallization and Data Collection**

592 $GEN1^{2\text{-}505\ D30N}$ and DNA (4w1010-1 GAATTCCGGATTAGGGATGC, 4w1010-2
593 GCATCCCTAAGCTCCATCGT, 4w1010-3 ACGATGGAGCCGCTAGGCTC, 4w1010-4
594 GAGCCTAGCGTCCGGAATTC) were mixed at a molar ratio of 2:1.1 at a final protein
595 concentration of 14 mg/ml including 1 mM $MgCl_2$ and co-crystallized by sitting drop vapor
596 diffusion. Drops were set up by mixing sample with mother liquor consisting of 100 mM
597 MES-NaOH pH 6.5 and 200 mM NaCl at a 2:1 ratio at room temperature. Crystals grew

598 within 2 days and several iterations of streak seeding were needed for obtaining diffraction
599 quality crystals. For data collection, crystals were stepwise soaked in 10%, 20% and 30%
600 (v/v) glycerol in 100 mM MES-NaOH pH 6.5, 200 mM NaCl and 5% PEG 8000 and flash-
601 frozen in liquid nitrogen. Diffraction data was collected at beamline PXII of the Swiss Light
602 Source (SLS, Villigen, Switzerland) at 100 K with a Pilatus 6M detector. In order to obtain
603 phase information, crystals were soaked for 10-30 min in 1 mM $[Ta_6Br_{12}]Br_2$, flash-frozen
604 and data was collected at the Ta L(III)-edge. In addition, seleno-methionine (SeMet)-
605 substituted protein was expressed in M9 media supplemented with SeMet, purified and
606 crystallized according to the protocol above and data was collected at the Se K-edge.

607 **Structure Determination and Refinement**

608 All data was processed with XDS (Table 1, Kabsch, 2010). HKL2MAP (Pape and Schneider,
609 2004) found 12 tantalum and 8 selenium positions, which were used in a combined MIRAS
610 strategy (multiple isomorphous replacement with anomalous scattering) in SHARP to
611 determine the structure of the GEN1-HJ complex. The obtained solvent-flattened
612 experimental map was used to build a model with PHENIX (Adams et al., 2010) combined
613 with manual building. The structure was then further refined by iterative rounds of manual
614 building in COOT (Emsley and Cowtan, 2004) and refinement with PHENIX. The structure
615 was visualized and analyzed in PYMOL (Delano, 2002). Electrostatic surface potentials
616 were calculated with PDB2PQR (Dolinsky et al., 2004) and APBS (Baker et al., 2001).

617 **Nuclease Assay**

618 All DNA substrates (Figure 5-figure supplement 1) were synthesized by Eurofins/MWG
619 (Ebersberg, Germany), resuspended in annealing buffer (20 mM Tris-HCl pH 8.0, 50 mM

620 NaCl, 0.1 mM EDTA), annealed by heating to 85˚C for 5 min and slow-cooling to room
621 temperature. Different amounts of GEN1 proteins (as indicated) were mixed with 40 nM
622 6FAM-labeled DNA substrates in 20 mM Tris-HCl pH 8.0, 50 ng/µl bovine serum albumin
623 (BSA) and 1 mM DTT. Reactions were initiated by adding 5 mM $MgCl_2$, incubated at 37°C
624 for 15 minutes and terminated by adding 15 mM EDTA, 0.3% SDS and further, DNA
625 substrates were deproteinized using 1 mg/ml proteinase K at 37°C for 15 minutes.
626 Products were separated by 8% 1x TBE native polyacrylamide gel electrophoresis, the
627 fluorescence signal detected with a Typhoon FLA 7000 phosphoimager (GE Healthcare),
628 quantified with IMAGEQUANT (GE Healthcare) and visualized by GNUPLOT (Williams et al.,
629 2015).

630 **Cruciform Plasmid Cleavage Assay**

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

639 **Sequence Alignments and Phylogenetic Analysis**

640 Sequences of GEN1 proteins from different organisms as well as all human chromodomain
641 proteins were aligned to the human GEN1 sequence using the programs HHPRED (Söding

642 et al., 2005), PSIBLAST and further by manual adjustments. Alignments were tested by
643 back-searches against RefSeq or HMM databases. A phylogenetic tree was calculated by the
644 program PHYML with 100 bootstraps using the alignment in Figure 3-figure supplement 2
645 and a BLOSUM62 substitution model. The tree was displayed with DENDROSCOPE (Huson
646 and Scornavacca, 2012).

647 **Histone Peptide Pull-down Assay**

648 The GEN1 chromodomain with a C-terminal His8-tag was immobilized on complete HisTag
649 Nickel resin and washed twice with binding buffer consisting of 20 mM Tris-HCl pH 7.5,
650 200 mM NaCl, 5% glycerol, 0.1 mM EDTA, 0.05% (v/v) Tween-20 and 2 mM TCEP. Peptide
651 mixtures containing 0.4 μM fluorescein labeled histone peptides were incubated with
652 beads at 4°C for 1 hour and washed twice with binding buffer. Immobilized proteins were
653 eluted with binding buffer supplemented with 300 mM imidazole and separated on 20%
654 SDS-PAGE. Fluorescein-labeled peptides were visualized by detecting the fluorescence
655 signal with a Typhoon FLA 7000 phosphoimager (GE Healthcare, Freiburg, Germany).

656 **Yeast Genetics and MMS Survival Assay in *Saccharomyces cerevisiae***

657 All yeast strains are based on W303 Rad5+ (see Figure 5-figure supplement 5 for a
658 complete list). *yen1Δ* or *yen1Δ mus81Δ* strains were transformed with an integrative
659 plasmid expressing mutant versions of *YEN1*. Freshly grown over-night cultures were
660 diluted to $1x10^7$ cells/ml. 5-fold serial dilutions were spotted on YPD plates with/without
661 MMS (methyl methanesulphonate, concentrations as indicated) and incubated for 2 days at
662 30°C. The expression of 3FLAG-tagged Yen1 constructs was verified by SDS-PAGE and

663 Western Blot analysis. Proteins were detected using a mouse monoclonal anti-FLAG M2-
664 peroxidase (HRP) antibody (Sigma-Aldrich, München, Germany).

## 665 Author Contributions

666 S.-H.L. and C.B. designed the study. S.-H.L., M.F.K. and C.B. purified proteins. S.-H.L.
667 performed crystallization experiments and activity assays. S.-H.L. and C.B. determined the
668 crystal structure. S.-H.L., B.H. and C.B. made sequence comparisons. L.N.P. and B.P.
669 performed yeast assays. S.-H.L. and C.B. prepared a draft and all authors were involved in
670 writing the manuscript.

## 671 Database Entry

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

## 674 Competing Interests Statement

675 The authors declare no competing interests.

## 676 Acknowledgements

677 We would like to thank Naoko Mizuno, Michael J. Taschner and Esben Lorentzen for many
678 scientific discussions and critical reading of the manuscript. We are grateful for the support
679 by the Department of Structural Cell Biology at the MPI of Biochemistry (MPIB),
680 particularly for the help with screening by the Crystallization Facility and Claire Basquin
681 for assistance with biophysical analysis. The Microchemistry Core Facility of the MPIB

682 provided mass spectrometry analysis. We would like to thank Jürg Müller for providing
683 histone tail peptides for binding tests and Stephen C. West for providing the plasmid
684 pIRbke8$^{mut}$ for functional assays. Furthermore, we would like to acknowledge the
685 professional assistance of the beamline staff at the Swiss Light Source, Villigen,
686 Switzerland, during data collection. The Max Planck Society for the Advancement of Science
687 supported this research through the Otto Hahn Program (B.P.) and the Max Planck
688 Research Group Leader Program (C.B.).

## 689 References

690 Adams, P.D., Afonine, P. V, Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung,
691 L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-
692 based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*,
693 213–221.

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

697 Ayoub, N., Jeyasekharan, A.D., Bernal, J. a, and Venkitaraman, A.R. (2008). HP1-beta
698 mobilization promotes chromatin changes that initiate the DNA damage response. Nature
699 *453*, 682–686.

700 Ayrapetov, M.K., Gursoy-Yuzugullu, O., Xu, C., Xu, Y., and Price, B.D. (2014). DNA double-
701 strand breaks promote methylation of histone H3 on lysine 9 and transient formation of
702 repressive chromatin. Proc Natl Acad Sci U S A *111*, 9169–9174.

703 Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of
704 nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A *98*,
705 10037–10041.

706 Bennett, R.J., and West, S.C. (1995a). RuvC protein resolves Holliday junctions via cleavage
707 of the continuous (noncrossover) strands. Proc Natl Acad Sci U S A *92*, 5635–5639.

708 Bennett, R.J., and West, S.C. (1995b). Structural analysis of the RuvC-Holliday junction
709 complex reveals an unfolded junction. J. Mol. Biol. *252*, 213–226.

710 Benson, F.E., and West, S.C. (1994). Substrate specificity of the Escherichia coli RuvC
711 protein. Resolution of three- and four-stranded recombination intermediates. J Biol Chem
712 *269*, 5195–5201.

713 Biertümpfel, C., Yang, W., and Suck, D. (2007). Crystal structure of T4 endonuclease VII
714 resolving a Holliday junction. Nature *449*, 616–620.

715 Blanco, M.G., and Matos, J. (2015). Hold your horSSEs: controlling structure-selective
716 endonucleases MUS81 and Yen1/GEN1. Front. Genet. *6*, 1–11.

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

720 Blanco, M.G., Matos, J., and West, S.C. (2014). Dual Control of Yen1 Nuclease Activity and
721 Cellular Localization by Cdk and Cdc14 Prevents Genome Instability. Mol. Cell *54*, 94–106.

722 Blus, B.J., Wiggins, K., and Khorasanizadeh, S. (2011). Epigenetic virtues of chromodomains.
723 Crit Rev Biochem Mol Biol *46*, 507–526.

724 Canzio, D., Chang, E.Y., Shankar, S., Kuchenbecker, K.M., Simon, M.D., Madhani, H.D.,
725 Narlikar, G.J., and Al-Sady, B. (2011). Chromodomain-mediated oligomerization of HP1
726 suggests a nucleosome-bridging mechanism for heterochromatin assembly. Mol Cell *41*,
727 67–81.

728 Castor, D., Nair, N., Déclais, A.-C., Lachaud, C., Toth, R., Macartney, T.J., Lilley, D.M.J., Arthur,
729 J.S.C., and Rouse, J. (2013). Cooperative Control of Holliday Junction Resolution and DNA
730 Repair by the SLX1 and MUS81-EME1 Nucleases. Mol. Cell *52*, 221–233.

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

734 Cejka, P., Plank, J.L., Dombrowski, C.C., and Kowalczykowski, S.C. (2012). Decatenation of
735 DNA by the S. cerevisiae Sgs1-Top3-Rmi1 and RPA Complex: A Mechanism for
736 Disentangling Chromosomes. Mol. Cell *47*, 886–896.

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

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

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

743 Cowieson, N.P., Partridge, J.F., Allshire, R.C., and McLaughlin, P.J. (2000). Dimerisation of a
744 chromo shadow domain and distinctions from the chromodomain as revealed by structural
745 analysis. Curr Biol *10*, 517–525.

746 Delano, W.L. (2002). The PyMOL Molecular Graphics System.

747 Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A. (2004). PDB2PQR: an
748 automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic
749 Acids Res *32*, W665–W667.

750 Eissenberg, J.C. (2012). Structural biology of the chromodomain: form and function. Gene
751 *496*, 69–78.

752 Eissler, C.L., Mazón, G., Powers, B.L., Savinov, S.N., Symington, L.S., and Hall, M.C. (2014).
753 The Cdk/Cdc14 Module Controls Activation of the Yen1 Holliday Junction Resolvase to
754 Promote Genome Stability. Mol. Cell *54*, 80–93.

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

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

760 Finger, L.D., Patel, N., Beddows, A., Ma, L., Exell, J.C., Jardine, E., Jones, A.C., and Grasby, J.A.
761 (2013). Observation of unpaired substrate DNA in the flap endonuclease-1 active site.
762 Nucleic Acids Res *41*, 9839–9847.

763 Fogg, J.M., and Lilley, D.M. (2000). Ensuring productive resolution by the junction-resolving
764 enzyme RuvC: large enhancement of the second-strand cleavage rate. Biochemistry *39*,
765 16125–16134.

766 García-Luis, J., Clemente-Blanco, A., Aragón, L., and Machín, F. (2014). Cdc14 targets the
767 Holliday junction resolvase Yen1 to the nucleus in early anaphase. Cell Cycle *13*, 1392–
768 1399.

769 Garner, E., Kim, Y., Lach, F.P., Kottemann, M.C., and Smogorzewska, A. (2013). Human GEN1
770 and the SLX4-Associated Nucleases MUS81 and SLX1 Are Essential for the Resolution of
771 Replication-Induced Holliday Junctions. Cell Rep. *5*, 207–215.

772 Giraud-Panis, M.J., and Lilley, D.M. (1997). Near-simultaneous DNA cleavage by the
773 subunits of the junction-resolving enzyme T4 endonuclease VII. EMBO J *16*, 2528–2534.

774 Górecka, K.M., Komorowska, W., and Nowotny, M. (2013). Crystal structure of RuvC
775 resolvase in complex with Holliday junction substrate. Nucleic Acids Res *41*, 9945–9955.

776 Grasby, J. a., Finger, L.D., Tsutakawa, S.E., Atack, J.M., and Tainer, J. a. (2012). Unpairing and
777 gating: Sequence-independent substrate recognition by FEN superfamily nucleases. Trends
778 Biochem. Sci. *37*, 74–84.

779 Hadden, J.M., Déclais, A.-C., Carr, S.B., Lilley, D.M.J., and Phillips, S.E. V (2007). The structural
780 basis of Holliday junction resolution by T7 endonuclease I. Nature *449*, 621–624.

781 Heyer, W. (2015). Regulation of Recombination and Genomic Maintenance. Cold Spring
782 Harb Perspect Biol *7*, a016501.

783 Holliday, R. (1964). A mechanism for gene conversion in fungi. Genet. Res. *5*, 282–304.

784 Holm, L., and Rosenström, P. (2010). Dali server: Conservation mapping in 3D. Nucleic
785 Acids Res. *38*, 545–549.

786 Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V., and Skrzypek, E. (2014).
787 PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. *43*, D512–
788 D520.

789 Huson, D.H., and Scornavacca, C. (2012). Dendroscope 3: An Interactive Tool for Rooted
790 Phylogenetic Trees and Networks. Syst. Biol. *61*, 1061–1067.

791 Interthal, H., and Heyer, W.D. (2000). MUS81 encodes a novel helix-hairpin-helix protein
792 involved in the response to UV- and methylation-induced DNA damage in Saccharomyces
793 cerevisiae. Mol. Gen. Genet. *263*, 812–827.

794 Ip, S.C.Y., Rass, U., Blanco, M.G., Flynn, H.R., Skehel, J.M., and West, S.C. (2008). Identification

795 of Holliday junction resolvases from humans and yeast. Nature *456*, 357–361.

796 Ira, G., Malkova, A., Liberi, G., Foiani, M., and Haber, J.E. (2003). Srs2 and Sgs1-Top3
797 Suppress Crossovers during Double-Strand Break Repair in Yeast. Cell *115*, 401–411.

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

801 Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. Acta
802 Crystallogr D Biol Crystallogr *66*, 133–144.

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

806 Lee, B., Nguyen, L.H., Barsky, D., Fernandes, M., and Wilson 3rd, D.M. (2002). Molecular
807 interactions of human Exo1 with DNA. Nucleic Acids Res *30*, 942–949.

808 Li, J., Li, Z., Ruan, J., Xu, C., Tong, Y., Pan, P.W., Tempel, W., Crombet, L., Min, J., and Zang, J.
809 (2011). Structural basis for specific binding of human MPP8 chromodomain to histone H3
810 methylated at lysine 9. PLoS One *6*, e25104.

811 Lieber, M.R. (1997). The FEN-1 family of structure-specific nucleases in eukaryotic DNA
812 replication, recombination and repair. Bioessays *19*, 233–240.

813 Lilley, D.M. (1985). The kinetic properties of cruciform extrusion are determined by DNA
814 base-sequence. Nucleic Acids Res *13*, 1443–1465.

815 Lilley, D.M., and White, M.F. (2001). The junction-resolving enzymes. Nat Rev Mol Cell Biol
816 *2*, 433–443.

817 Matos, J., and West, S.C. (2014). Holliday junction resolution: Regulation in space and time.
818 DNA Repair (Amst). *19*, 176–181.

819 Matos, J., Blanco, M.G., Maslen, S., Skehel, J.M., and West, S.C. (2011). Regulatory control of
820 the resolution of DNA recombination intermediates during meiosis and mitosis. Cell *147*,
821 158–172.

822 Miętus, M., Nowak, E., Jaciuk, M., Kustosz, P., Studnicka, J., and Nowotny, M. (2014). Crystal
823 structure of the catalytic core of Rad2: insights into the mechanism of substrate binding.
824 Nucleic Acids Res *42*, 10762–10775.

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

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

830 O'Shaughnessy, A., and Hendrich, B. (2013). CHD4 in the DNA-damage response and cell
831 cycle progression: not so NuRDy now. Biochem. Soc. Trans. *41*, 777–782.

832 Orans, J., McSweeney, E.A., Iyer, R.R., Hast, M.A., Hellinga, H.W., Modrich, P., and Beese, L.S.
833 (2011). Structures of human exonuclease 1 DNA complexes suggest a unified mechanism
834 for nuclease family. Cell *145*, 212–223.

835 Pape, T., and Schneider, T.R. (2004). HKL2MAP: a graphical user interface for
836 macromolecular phasing with SHELX programs. J. Appl. Crystallogr. *37*, 843–844.

837 Patel, N., Atack, J.M., Finger, L.D., Exell, J.C., Thompson, P., Tsutakawa, S., Tainer, J.A.,
838 Williams, D.M., and Grasby, J.A. (2012). Flap endonucleases pass 5'-flaps through a flexible
839 arch using a disorder-thread-order mechanism to confer specificity for free 5'-ends.
840 Nucleic Acids Res *40*, 4507–4519.

841 Petronczki, M., Siomos, M.F., and Nasmyth, K. (2003). Un Ménage à Quatre. Cell *112*, 423–
842 440.

843 Pottmeyer, S., and Kemper, B. (1992). T4 endonuclease VII resolves cruciform DNA with
844 nick and counter-nick and its activity is directed by local nucleotide sequence. J Mol Biol
845 *223*, 607–615.

846 Price, B.D., and D'Andrea, A.D. (2013). Chromatin remodeling at DNA double-strand breaks.
847 Cell *152*, 1344–1354.

848 Putnam, C.D., Hayes, T.K., and Kolodner, R.D. (2009). Specific pathways prevent
849 duplication-mediated genome rearrangements. Nature *460*, 984–989.

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

853 Sakurai, S., Kitano, K., Yamaguchi, H., Hamada, K., Okada, K., Fukuda, K., Uchida, M., Ohtsuka,
854 E., Morioka, H., and Hakoshima, T. (2005). Structural basis for recruitment of human flap
855 endonuclease 1 to PCNA. EMBO J. *24*, 683–693.

856 Sarbajna, S., and West, S.C. (2014). Holliday junction processing enzymes as guardians of
857 genome stability. Trends Biochem Sci *39*, 409–419.

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

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

862 Shah, R., Cosstick, R., and West, S.C. (1997). The RuvC protein dimer resolves Holliday
863 junctions by a dual incision mechanism that involves base-specific contacts. EMBO J *16*,
864 1464–1472.

865 Smothers, J.F., and Henikoff, S. (2000). The HP1 chromo shadow domain binds a consensus
866 peptide pentamer. Curr. Biol. *10*, 27–30.

867 Söding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein
868 homology detection and structure prediction. Nucleic Acids Res *33*, W244–W248.

869 Stanley, F.K.T., Moore, S., and Goodarzi, A. a. (2013). CHD chromatin remodelling enzymes
870 and the DNA damage response. Mutat. Res. - Fundam. Mol. Mech. Mutagen. *750*, 31–44.

871 Svendsen, J.M., and Harper, J.W. (2010). GEN1/Yen1 and the SLX4 complex: Solutions to the
872 problem of Holliday junction resolution. Genes Dev *24*, 521–536.

873 Svendsen, J.M., Smogorzewska, A., Sowa, M.E., O'Connell, B.C., Gygi, S.P., Elledge, S.J., and
874 Harper, J.W. (2009). Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase

875 and is required for DNA repair. Cell *138*, 63–77.

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

878 Tomlinson, C.G., Atack, J.M., Chapados, B., Tainer, J.A., and Grasby, J.A. (2010). Substrate
879 recognition and catalysis by flap endonucleases and related enzymes. Biochem Soc Trans
880 *38*, 433–437.

881 Tsutakawa, S.E., Classen, S., Chapados, B.R., Arvai, A.S., Finger, L.D., Guenther, G., Tomlinson,
882 C.G., Thompson, P., Sarker, A.H., Shen, B., et al. (2011). Human flap endonuclease structures,
883 DNA double-base flipping, and a unified understanding of the FEN1 superfamily. Cell *145*,
884 198–211.

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

888 Wakasugi, M., Reardon, J.T., and Sancar, A. (1997). The Non-catalytic function of XPG
889 protein during dual incision in human nucleotide excision repair. J. Biol. Chem. *272*, 16030–
890 16034.

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

893 West, S.C., Blanco, M.G., Chan, Y.W., Matos, J., Sarbajna, S., and Wyatt, H.D.M. (2015).
894 Resolution of Recombination Intermediates: Mechanisms and Regulation. Cold Spring
895 Harb. Symp. Quant. Biol. *LXXX*.

896 Williams, T., Kelley, C., Campbell, J., Cunningham, R., Denholm, D., Elber, G., Fearick, R.,
897 Grammes, C., Hart, L., Hecking, L., et al. (2015). gnuplot 5.0. *0*.

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

900 Wyatt, H.D.M., Sarbajna, S., Matos, J., and West, S.C. (2013). Coordinated Actions of SLX1-
901 SLX4 and MUS81-EME1 for Holliday Junction Resolution in Human Cells. Mol. Cell *52*, 234–
902 247.

903 Yang, W. (2011). Nucleases: diversity of structure, function and mechanism. Q Rev Biophys
904 *44*, 1–93.

905 Yap, K.L., and Zhou, M.-M. (2011). Structure and mechanisms of lysine methylation
906 recognition by the chromodomain in gene transcription. Biochemistry *50*, 1966–1980.

907 Ying, S., Minocherhomji, S., Chan, K.L., Palmai-Pallag, T., Chu, W.K., Wass, T., Mankouri, H.W.,
908 Liu, Y., and Hickson, I.D. (2013). MUS81 promotes common fragile site expression. Nat. Cell
909 Biol. *15*, 1001–1007.

910

911

912

913 **Table 1**

914 **Data collection and refinement statistics.**

| Data Set | G505-4w006 native | G505-4w006 Ta peak | G505-4w006 SeMet 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/σI** | 13.39 (1.42) | 27.49 (5.83) | 16.58 (3.82) |
| Completeness (%)* | 99.3 (98.5) | 99.6 (97.3) | 97.3 (83.3) |
| Redundancy* | 3.2 | 10.2 | 5.1 |
| *Rsym* (%)* | 6.2 (82.2) | 7.7 (42.2) | 6.9 (43.4) |
| **Refinement Statistics** | | | |
| Resolution (Å) | 75-3.0 | | |
| Number of Reflections | 67667 | | |
| $R_{work}/R_{free}$ | 0.219/0.259 | | |
| Number of Atoms | | | |
| Protein | 6246 | | |
| DNA | 1609 | | |
| Water/Solutes | 20 | | |
| B-factors | | | |
| Protein | 112.7 | | |
| DNA | 142.4 | | |
| Water/Solutes | 200.0 | | |
| R.M.S Deviations | | | |
| Bond lengths (Å) | 0.005 | | |
| Bond Angles (°) | 0.813 | | |
| Ramachandran Plot | | | |
| Preferred | 724 (94.8 %) | | |
| Allowed | 40 (5.2%) | | |

915 *Values for the highest resolution shell are shown in parenthesis
916
917

918

919 ## Figure 1

920 **Architecture of human GEN1.** (**A**) Domain architecture of human GEN1. The structurally
921 unknown regulatory domain (residues 465-908) is shown with dotted lines. (**B**) Overview
922 of the catalytic core of GEN1 in complex with HJ DNA. The protein resembles the shape of a
923 downwards-pointing right hand with helix α6 as the thumb. The protein is depicted in half
924 transparent surface representation with secondary structure elements underneath. The
925 DNA is shown in ladder representation with individual strands in different colors. The
926 coloring of GEN1 follows domain boundaries: intertwining XPG-N and XPG-I in green, 5'->3'
927 exonuclease C-terminal domain (EXO) in blue, chromodomain in pink, unassigned regions
928 in gray. Active site residues (E134, E136, D155, D157) are highlighted in orange. (**C**)
929 Electrostatic surface potential of GEN1. The coloring follows the potential from -5 (red) to
930 +5 kT/e (blue). The DNA-binding interfaces and the position of the hydrophobic wedge are
931 marked in yellow. (**D**) Secondary structure elements of the catalytic core of GEN1 in
932 cartoon representation with the same colors as before. Dotted lines represent parts that
933 are not resolved in the crystal structure. The numbering follows a unified scheme for the
934 Rad2/XPG family (compare Figure 2) for α-helices, β-sheets and $3_{10}$-helices (η). (**E**)
935 Experimental electron density map (autoSHARP, solvent flattened, contoured at 1σ) drawn
936 around the HJ in the GEN1 complex. The DNA model is shown in ball-stick representation
937 with carbon atoms of individual strands in different colors (yellow, light blue, magenta,
938 green) and oxygen atoms in red, phosphor atoms in orange, nitrogen atoms in dark blue.
939 (**F**) Structural comparison of Rad2/XPG family nucleases. Proteins are shown in a
940 simplified surface representation with important structural elements in cartoon
941 representation and DNA in ladder representation. The color scheme is the same as in B.
942 **Figure 1-figure supplement 1** shows the content of the asymmetric unit.

943 ## Figure 2

944 **Alignment of the nuclease cores of Rad2/XPG-family proteins**. The alignment is based
945 on known crystal structures: human GEN1 (PDB xxxx, this study), yeast Rad2 (PDB 4q0w),
946 human FEN1 (PDB 3q8k), human EXO1 (3qe9). Secondary structure elements are depicted

947 on top of the sequence with dark blue bars for α-helices, light blue bars for $3_{10}$-helices and
948 green arrows for β-sheets. The numbering follows a unified scheme for the superfamily.
949 Functional elements are labeled and described in the main text. Sequences are colored by
950 similarity (BLOSUM62 score) and active site residues are marked in red. Mutations
951 analyzed in this study are marked with an orange triangle and DNA contacts found in the
952 human GEN1–HJ structure have a dark green dot. Disordered or missing parts in the
953 structures are labeled in small letters or with "x".

954 **Figure 3**

955 **Chromodomain comparison.** (**A**) Sequence alignment of GEN1 chromodomains from
956 different organisms: hsGEN1 (*Homo sapiens*), clGEN1 (*Canis lupus*), mmGEN1 (*Mus*
957 *musculus*), drGEN1 (*Danio rerio*), atGEN1/2 (*Arabidopsis thaliana*), cgGEN1 (*Crassostrea*
958 *gigas*), scYEN1 (*Saccharomyces cerevisiae*). The presence of a chromodomain is conserved
959 from yeast to human with *Caenorhabditis elegans* as an exception. Secondary structure
960 elements of the GEN1 chromodomain are shown on top. The sequence coloring is based on
961 a similarity matrix (BLOSUM62). The corresponding positions of the DNA-interaction site
962 in human GEN1 is marked with a red box and residues of the aromatic cage are highlighted
963 with a yellow box. (**B**) GEN1 has a canonical chromodomain fold of three antiparallel beta-
964 sheets packed against an α-helix. (**C**) The arrangement of the aromatic cage in GEN1 is
965 comparable to other chromodomains but less aromatic and slightly larger. (**D**) The
966 superposition of different chromodomains places cognate binding peptides of hsMPP8 and
967 mmCBX7 (and others) into the aromatic cage. (**E**) The aromatic cage of GEN1 is closed by
968 helix α15. Panels B-D show the chromodomains of hsGEN1 (pink, PDB xxxx), hsCBX3 (gray,
969 PDB 3kup) hsSUV39H1 (green, PDB 3mts), hsMPP8 (yellow, PDB 3lwe), dmHP1a (orange,
970 chromo shadow PDB 3p7j), dmRHINO (cyan, PDB 4quc/3r93), mmCBX7 (light blue, PDB
971 4x3s; compare **Figure 3-figure supplement 1**). (**F**) Phylogenetic tree of all known human
972 chromodomains. GEN1 is distantly related to the CBX chromo-shadow domains and CDY
973 chromodomains. The corresponding alignment for calculating the phylogenetic tree is
974 shown in **Figure 3-figure supplement 2**. GEN1 is colored in black, chromobox (CBX)
975 proteins are colored in red, interspersed by SUV39H histone acetylases (orange) and
976 chromodomain Y-linked (CDY) proteins (yellow). Chromo-barrel domain proteins are
977 colored in green and chromodomain-helicase DNA-binding (CHD) proteins are in blue.

978 Chromodomains and chromo-shadow domains from the same protein are labeled with 1
979 and 2, respectively. Stable branches with boostrap values equal or higher than 0.8 are
980 marked with a black dot. The binding of the GEN1 chromodomain to a set of histone
981 peptides was tested but no interaction was detected (**Figure 3-figure supplement 3** and
982 **Figure 3-figure supplement 4**).

983 ## Figure 4

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

998 ## Figure 5

999 **Functional analysis of GEN1.** (**A**) Nuclease activity of GEN1 with HJ and 5'flap DNA. 40 nM
1000 5' 6FAM-labeled substrates were mixed with indicated amounts of GEN1. Reactions were
1001 carried out at 37°C for 15 minutes, products were separated by native PAGE and analyzed
1002 with a phosphoimager. **Figure 5-figure supplement 1** gives the sequences of DNA oligos
1003 used in biochemical assays and **Figure 5 source data 1** shows activity measurements. (**B**)
1004 Quantification of nuclease assays of wild type GEN1 and variants with mutated residues
1005 located at the protein-DNA interfaces. Percentage of cleavage was plotted against the
1006 enzyme concentration. Error bars depict the standard deviation calculated from at least
1007 three independent experiments. **Figure 5-figure supplement 2** shows representative gels

1008 from the PAGE analysis. (**C**) Quantification of nuclease assays of wild type GEN1 and
1009 variants with mutated residues located at the chromodomain. Error bars depict the
1010 standard deviation calculated from at least three independent experiments. **Figure 5-**
1011 **figure supplement 3** shows representative gels from the PAGE analysis. (**D**) GEN1
1012 mutations used in this study. Locations of human GEN1 mutations used in biochemical
1013 assays and corresponding residues in yeast MMS survival assays are highlighted in red.
1014 Active site residues E134, E136, D155, D157 are marked in turquoise. (**E**) Schematic of the
1015 cruciform plasmid cleavage assay. A cruciform structure can be formed in plasmid
1016 pIRbke8$^{mut}$, which harbors an inverted-repeat sequence and is stabilized by negative
1017 supercoiling. Introducing two cuts across the junction point within the lifetime of the
1018 resolvase-junction complex yields linear products whereas sequential cleavage generates
1019 nicked products and the relaxed plasmid cannot be a substrate for the next cleavage. (**F**)
1020 Cruciform plasmid cleavage assay with different GEN1 variants. Plasmid pIRbke8$^{mut}$ was
1021 treated with 256 nM GEN1 each and reactions were carried out at 37°C for 15 minutes.
1022 Supercoiled, linear and nicked plasmids were separated by native agarose gel
1023 electrophoresis and visualized with SYBR safe under UV light. (**G**) MMS survival assays
1024 with yeast y*en1* variants. The survival of *yen1* mutants was tested under a *yen1Δ mus81Δ*
1025 background with indicated amounts of MMS. The top part shows mutations at GEN1-DNA
1026 interfaces and the bottom part mutations at the chromodomain (compare **Figure 5-figure**
1027 **supplement 4** for all controls and expression tests). **Figure 5-figure supplement 5** gives
1028 a list of all yeast strains.

## 1029 Figure 6

1030 **Substrate recognition features of GEN1.** (**A**) Superposition of the protein part of the
1031 FEN1-DNA complex (PDB 3q8k, protein in gray, DNA in black) onto the GEN1-HJ complex
1032 (protein in green and the DNA strands in different colors). The FEN1-DNA aligns with the
1033 same register as the GEN1-DNA at the upstream interface. (**B**) Model for the recognition of
1034 a 5' flap substrate by GEN1. The DNA was extended using the superimposition from A.
1035 Homology modeling suggests an additional helix α4 (disordered residues 79-92) forming
1036 an arch with helix α6. The protein is shown in a simplified surface representation with the
1037 same colors as in Figure 1 and structural elements are highlighted. The insert shows a
1038 zoomed in view of the hydrophobic wedge with the modeled FEN1-DNA in gray. (**C**) Model

1039 for the dimerization of GEN1 upon binding to a HJ substrate based on the 5' flap model in B.
1040 The monomers interlock via both arches (α4-α6) and the hydrophobic wedges (α2-α3)
1041 contact each other. (**D**) Structure of the *Thermus thermophilus* RuvC-HJ complex (PDB
1042 4ld0). (**E**) Structure of the T4 endonuclease VII-HJ complex (PDB 2qnc). (**F**) Structure of the
1043 T7 endonuclease I-HJ complex (PDB 2pfj). Individual monomers are in surface
1044 representation, colored in light blue and beige, respectively. DNA strands are shown as
1045 ladders in different colors.

## 1046 Figure 1-figure supplement 1

1047 **Content of the asymmetric unit of the GEN1-HJ crystal.** One protein monomer is shown
1048 in surface representation with secondary structure cartoons underneath, the other one
1049 only in cartoon representation with α-helices as cylinders and β-strands as arrows. The HJ
1050 bridges between two protein monomers in the asymmetric unit. The active sites are
1051 labeled with a turquoise ball.

## 1052 Figure 3- figure supplement 1

1053 **Proteins found in a DALI search.** Top hits found in a DALI search for protein structure
1054 comparison with the human GEN1 chromodomain (residues 390-464) against the Protein
1055 Data Bank. The most similar unique chromodomains are listed.

## 1056 Figure 3- figure supplement 2

1057 **Sequence alignment of all known human chromodomains.** The alignment was used to
1058 calculate the phylogenetic tree in Figure 3F. Colors follow the CLUSTAL X coloring scheme.

## 1059 Figure 3- figure supplement 3

1060 **N-terminally fluorescein-labeled peptides used for chromodomain binding assays.**

## 1061 Figure 3- figure supplement 4

1062 **Histone peptide pull-down assay.** Nickel resin-immobilized GEN1 chromodomain was
1063 incubated with mixtures of fluorescein-labeled histone peptides, washed, bound peptides
1064 eluted and separated by 20% SDS-PAGE. Mix 1 and 2 did not show any binding and non-

1065 specific binding to the resin was found with Mix 3. The smearing of the bands is due to the
1066 small size of the peptides (~1.5 kDa). I, C and E represent input, resin control and elution,
1067 respectively. Mix 1: H3K9, H3K9me1, H3K9me2 and H3K9me3. Mix 2: H3K27, H3K27me1,
1068 H3K27me2 and H3K27me3. Mix 3: H3K36me1, H3K36me2, H3K36me3 and H3K36Ac.

## 1069 Figure 5- figure supplement 1

1070 **Oligonucleotides used in biochemical assays.** Four-way junctions were prepared by
1071 annealing CB209, CB210, CB211, CB212. 5' flaps were prepared by annealing CB209,
1072 CB212, CB218. The annealing protocol is described in Material and Methods.

## 1073 Figure 5- figure supplement 2

1074 **DNA cleavage assays of different GEN1 mutations.** All $GEN1^{2-505}$ mutations were
1075 generated by site-directed mutagenesis and purified with the same procedure.
1076 Experiments were repeated three times and a representative gel picture is shown for each
1077 protein variant in Figure 5.

## 1078 Figure 5- figure supplement 3

1079 **DNA cleavage assays of different GEN1 fragments.** (**A**) 5' 6FAM labeled four-way
1080 junction or 5'flap DNA (40 nM) were mixed with varying concentrations of GEN1
1081 truncations (0.25, 0.5, 1, 2, 4, 8, 16, 32, 64, 128, 256 nM, respectively). (**B**) Quantification of
1082 activity assays.

## 1083 Figure 5- figure supplement 4

1084 **MMS survival assays with yeast *yen1* mutants.** The survival of *yen1* mutants was tested
1085 in a *yen1Δ* or *yen1Δ mus81Δ* background with indicated amounts of MMS. (compare Figure
1086 5 and Figure 5-figure supplement 5). Mus81 overlaps with Yen1 functionally, therefore
1087 *yen1Δ* knock-out strains are fully viable even in the presence of MMS and hypersensitivity
1088 is only seen in the double knock-out. (**A**) Mutations in the chromodomain. (**B**) Mutations at
1089 protein-DNA interfaces. (**C**) *Yen1* truncations and chromodomain deletion. (**D**) Protein
1090 expression test (Western Blot analysis) of 3FLAG tagged Yen1 variants. Asterisk denotes a
1091 cross-reactive band.

1092 ## Figure 5- figure supplement 5

1093 **Yeast strains used for MMS survival assays.**

1094 ## Figure 5-source data 1

1095 ***In vitro* activity measurements of different GEN1$^{2-505}$ variants.**

1096 ## List of Files

1097 **Figure_1.pdf** Architecture of human GEN1

1098 **Figure_2.pdf** Alignment of the nuclease cores of Rad2/XPG-family proteins

1099 **Figure_3.pdf** Chromodomain comparison

1100 **Figure_4.pdf** DNA interactions in the GEN1-DNA complex

1101 **Figure_5.pdf** Functional analysis of GEN1

1102 **Figure_6.pdf** Substrate recognition features of GEN1

1103 **Figure_1_supplement_1.pdf** Content of the asymmetric unit of the GEN1-HJ crystal

1104 **Figure_3_supplement_1.docx** Proteins found in a DALI search

1105 **Figure_3_supplement_2.pdf** Sequence alignment of all known human chromodomains

1106 **Figure_3_supplement_3.docx** N-terminally fluorescein-labeled peptides used for
1107 chromodomain binding assays

1108 **Figure_3_supplement_4.pdf** Histone peptide pull-down assay

1109 **Figure_5_supplement_1.docx** Oligonucleotides used in biochemical assays

1110 **Figure_5_supplement_2.pdf** DNA cleavage assays of different GEN1 mutations

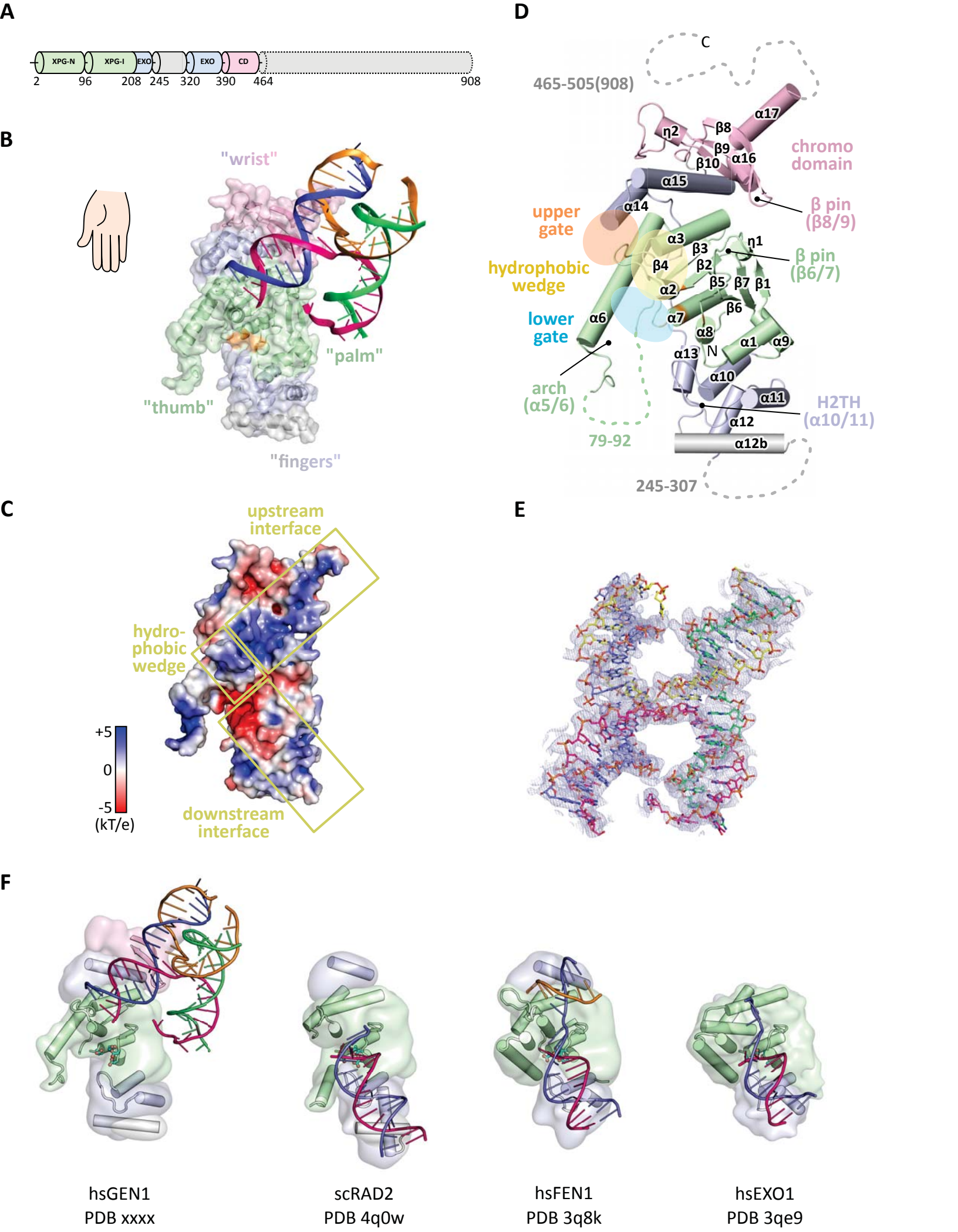
1111 **Figure_5_supplement_3.pdf** DNA cleavage assays of different GEN1 fragments

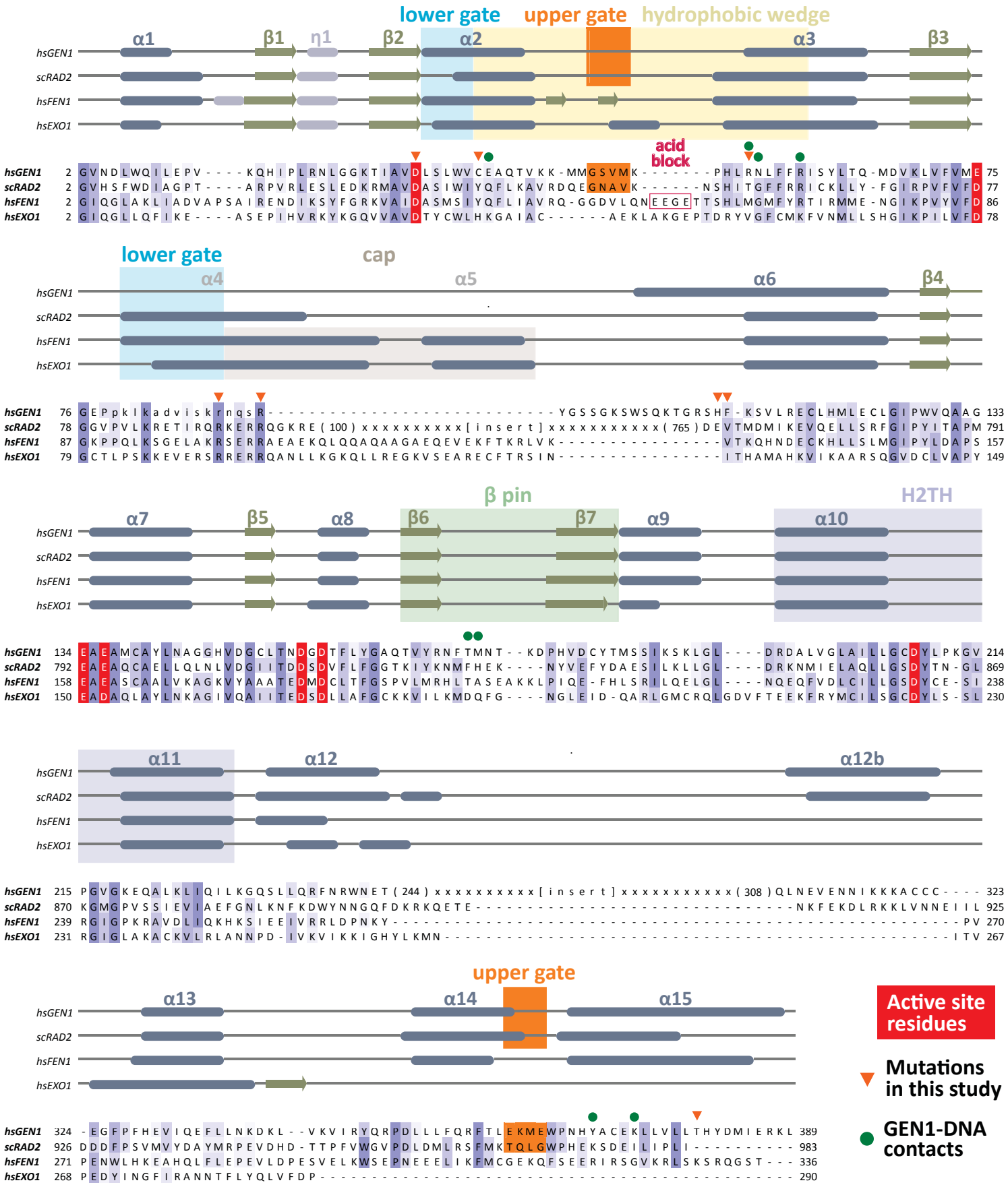
1112 **Figure_5_supplement_4.pdf** MMS survival assays with yeast *yen1* mutants

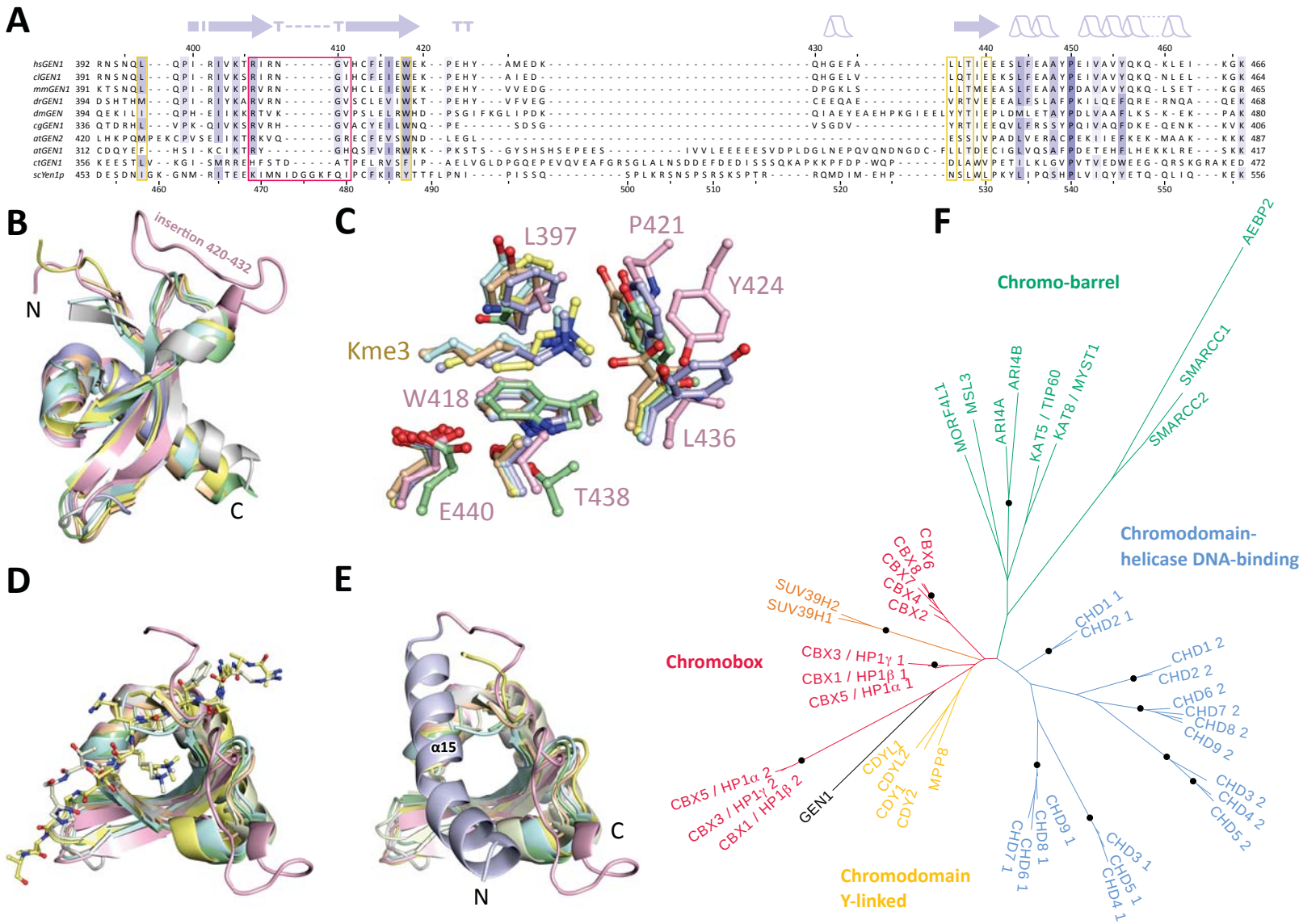
1113 **Figure_5_supplement_5.docx** Yeast strains used for MMS survival assays

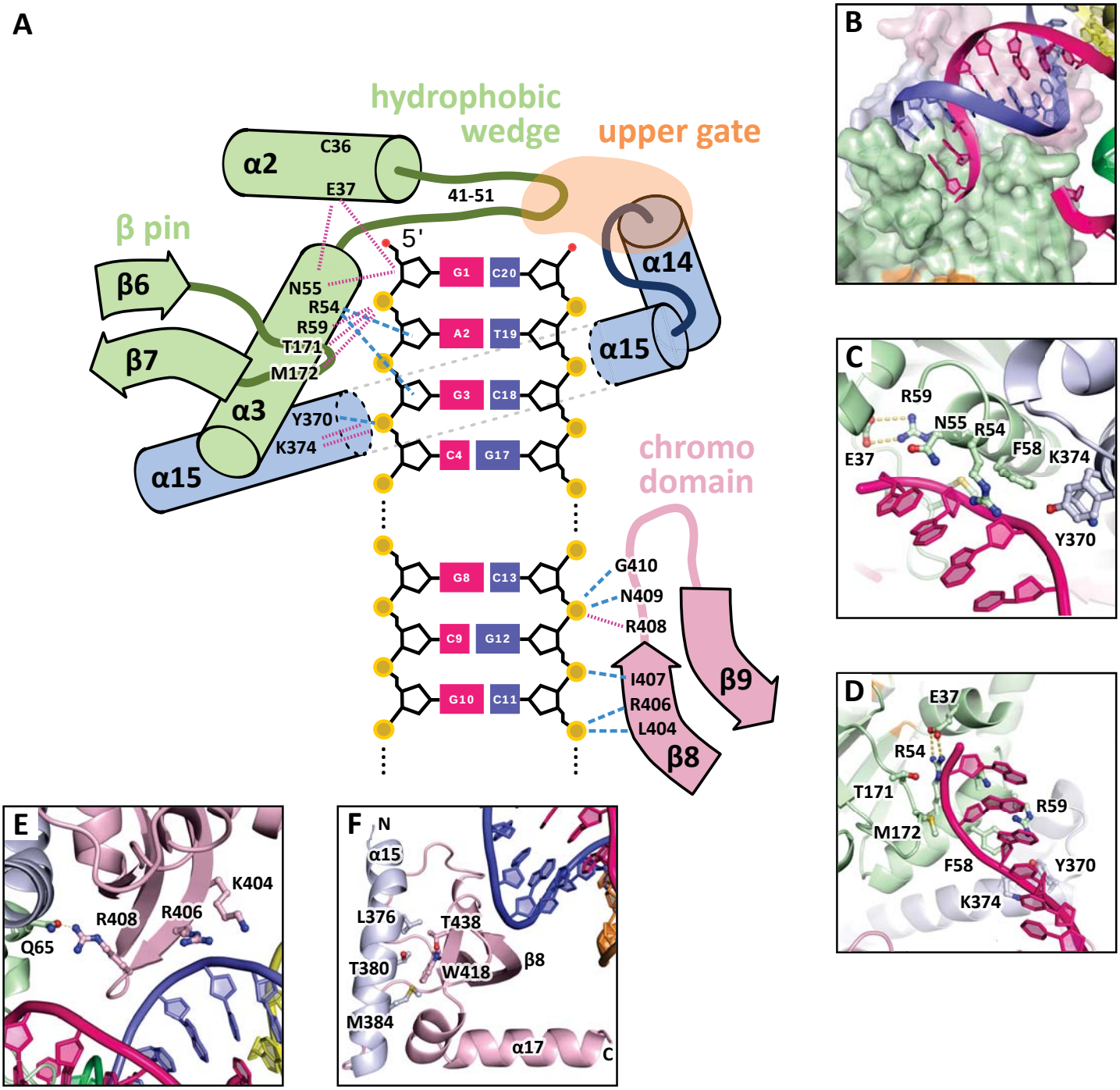
1114 **Figure_5_source_data_1.xlsx** *In vitro* activity measurements of different GEN1$^{2-505}$
1115 variants.

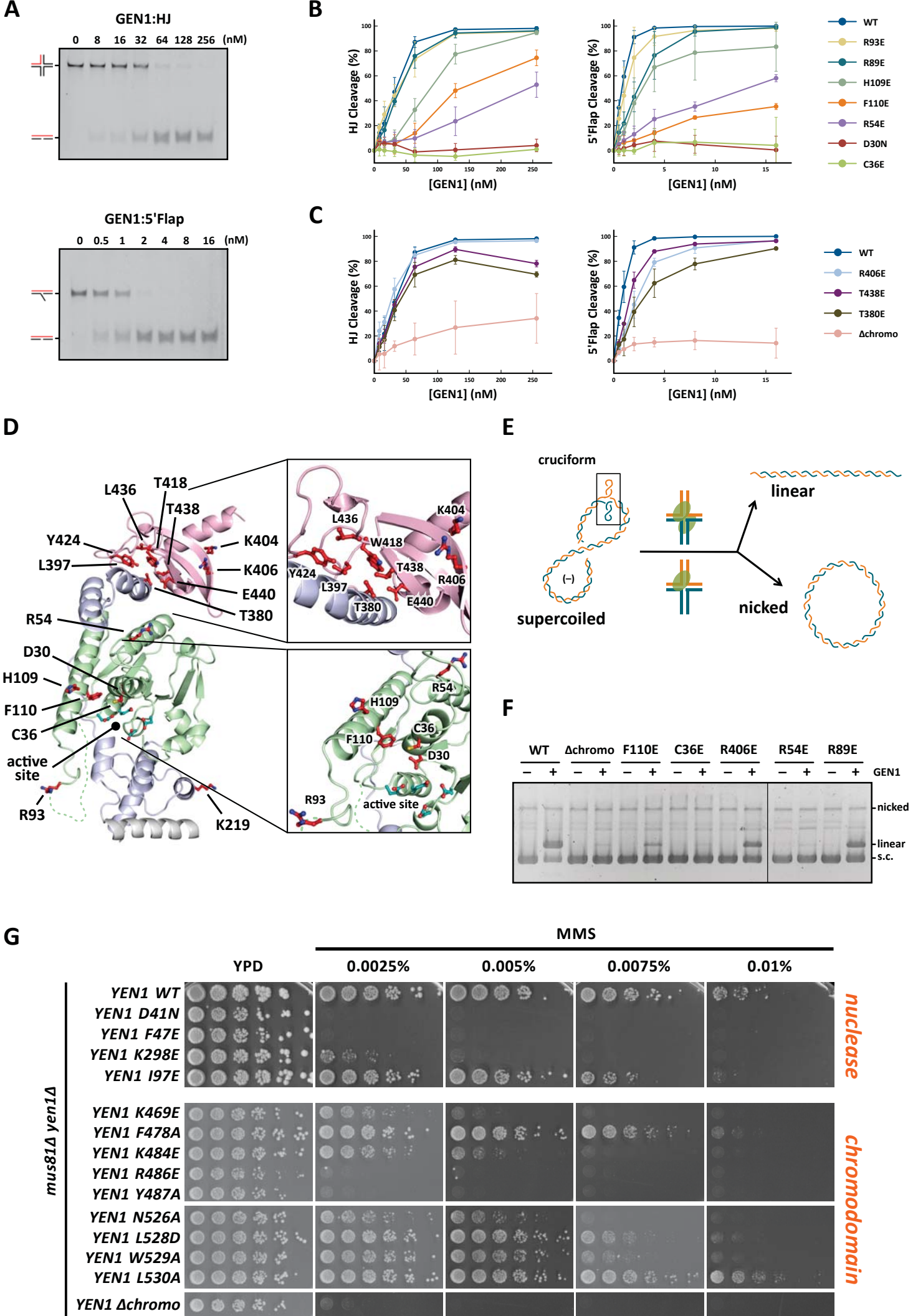
1116

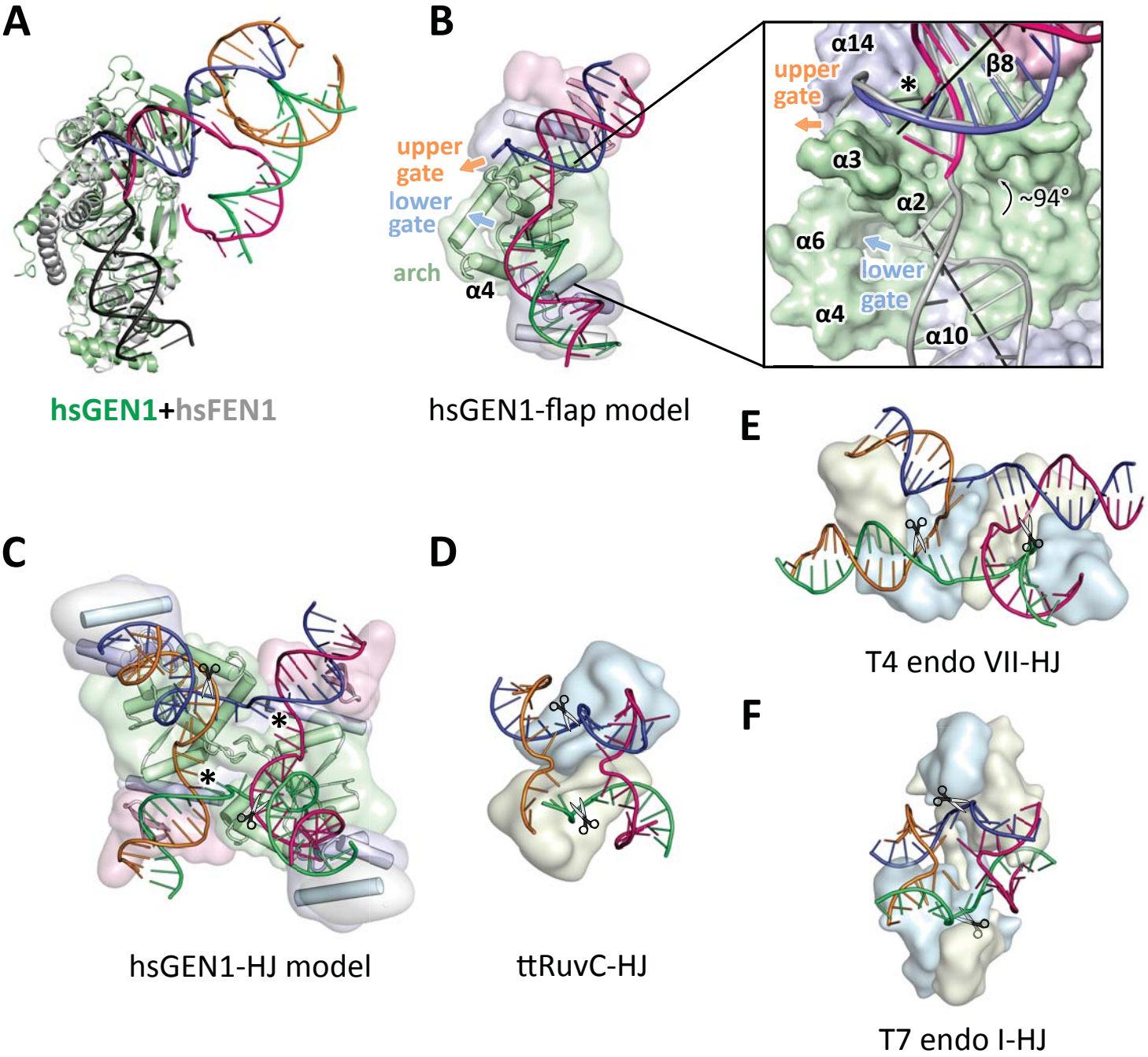

A
XPG-N
XPG-I
EXO
EXO
CD
2
96
208
245
320
390
464
908
B
"wrist"
"palm"
"thumb"
"fingers"
C
upstream interface
hydro-phobic wedge
downstream interface
+5
0
-5
(kT/e)
D
C
465-505(908)
α17
η2
β8
β9
β10
α16
chromo domain
α15
α14
upper gate
β pin (β8/9)
α3
β3
η1
β4
β2
β pin (β6/7)
hydrophobic wedge
α2
β5
β7
β1
β6
lower gate
α6
α7
α8
α1
α9
α13
N
α10
arch (α5/6)
α11
H2TH (α10/11)
α12
79-92
α12b
245-307
E
F
hsGEN1
PDB xxxx
scRAD2
PDB 4q0w
hsFEN1
PDB 3q8k
hsEXO1
PDB 3qe9


lower gate
upper gate
hydrophobic wedge
α1
β1
η1
β2
α2
α3
β3
hsGEN1
scRAD2
hsFEN1
hsEXO1
acid block
hsGEN1 2 GVNDLWQILEPV----KQHIPLRNLGGKTIAVDLSLWVCEAQTVKK-MMGSVMK------PHLRNLFFRISYLTQ-MDVKLVFVME 75
scRAD2 2 GVHSFWDIAGPT----ARPVRLESLEDKRMAVDASIWIYQFLKAVRDQEGNAVK-----NSHITGFFRRICKLLY-FGIRPVFVFD 77
hsFEN1 2 GIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQ-GGDVLQNEEGETTSHLMGMFYRTIRMME-NGIKPVYVFD 86
hsEXO1 2 GIQGLLQFIKE----ASEPIHVRKYKGQVVAVDTYCWLHKGAIAC------AEKLAKGEPTDRYVGFCMKFVNMLLSHGIKPILVFD 78
lower gate
cap
α4
α5
α6
β4
hsGEN1 76 GEPpkIkadviskrnqsR----------YGSSGKSWSQKTGRSHF-KSVLRECLHMLECLGIPWVQAAG 133
scRAD2 78 GGVPVLKRETIRQRKERRQGKRE(100)xxxxxxxxxx[insert]xxxxxxxxxxx(765)DEVTMDMIKEVQELLSRFGIPYITAPM 791
hsFEN1 87 GKPPQLKSGELAKRSERRAEAEKQLQQAQAAGAEQEVEKFTKRLVK----------VTKQHNDECKHLLSLMGIPYLDAPS 157
hsEXO1 79 GCTLPSKKEVERSRRERRQANLLKGKQLLREGKVSEARECFTRSIN----------ITHAMAHKVIKAARSQGVDCLVAPY 149
β pin
H2TH
α7
β5
α8
β6
β7
α9
α10
hsGEN1 134 EAEAMCAYLNAGGHVDGCLTNDGDTFLYGAQTVYRNFTMNT--KDPHVDCYTMSSIKSKLGL----DRDALVGLAILLGCDYLPKGV 214
scRAD2 792 EAEAQCAELLQLNLVDGIITDDSDVFLFGGTKIYKNMFHEK----NYVEFYDAESILKLLGL----DRKNMIELAQLLGSDYTN-GL 869
hsFEN1 158 EAEASCAALVKAGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLPIQE-FHLSRILQELGL----NQEQFVDLCILLGSDYCE-SI 238
hsEXO1 150 EADAQLAYLNKAGIVQAIITEDSDLLAFGCKKVILKMDQFG----NGLEID-QARLGMCRQLGDVFTEEKFRYMCILSGCDYLS-SL 230
α11
α12
α12b
hsGEN1 215 PGVGKEQALKLIQILKGQSLLQRFNRWNET(244)xxxxxxxxxx[insert]xxxxxxxxxxx(308)QLNEVENNIKKKACCC---- 323
scRAD2 870 KGMGPVSSIEVIAEFGNLKNFKDWYNNGQFDKRKQETE----------NKFEKDLRKKLVNNEIIL 925
hsFEN1 239 RGIGPKRAVDLIQKHKSIEEIVRRLDPNKY----------PV 270
hsEXO1 231 RGIGLAKACKVLRLANNPD-IVKVIKKIGHYLKMN----------ITV 267
upper gate
α13
α14
α15
hsGEN1 324 -EGFPFHEVIQEFLLNKDKL--VKVIRYQRPDLLLFQRFTLEKMEWPNHYACEKLLVLLTHYDMIERKL 389
scRAD2 926 DDDFPSVMVYDAYMRPEVDHD-TTPFVWGVPDLDMLRSFMKTQLGWPHEKSDEILIPLI---------- 983
hsFEN1 271 PENWLHKEAHQLFLEPEVLDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGVKRLSKSRQGST--- 336
hsEXO1 268 PEDYINGFIRANNTFLYQLVFDP---------- 290
Active site residues
Mutations in this study
GEN1-DNA contacts


A
B
insertion 420-432
N
C
C
L397
P421
Y424
Kme3
W418
L436
E440
T438
D
E
α15
C
N
F
AEBP2
Chromo-barrel
SMARCC1
SMARCC2
MORF4L1
MSL3
ARI4A
ARI4B
KAT5 / TIP60
KAT8 / MYST1
Chromodomain-
helicase DNA-binding
CBX6
CBX8
CBX7
CBX4
CBX2
SUV39H2
SUV39H1
Chromobox
CBX3 / HP1γ 1
CBX1 / HP1β 1
CBX5 / HP1α 1
CBX5 / HP1α 2
CBX3 / HP1γ 2
CBX1 / HP1β 2
GEN1
CDYL1
CDYL2
CDY1
CDY2
MPP8
CHD1 1
CHD2 1
CHD1 2
CHD2 2
CHD6 2
CHD7 2
CHD8 2
CHD9 2
CHD3 2
CHD4 2
CHD5 2
CHD9 1
CHD8 1
CHD6 1
CHD7 1
CHD3 1
CHD5 1
CHD4 1
Chromodomain
Y-linked


A
hydrophobic wedge
upper gate
α2
C36
E37
41-51
β pin
5'
β6
β7
N55
R54
R59
T171
M172
α3
Y370
K374
α15
α14
α15
G1
C20
A2
T19
G3
C18
C4
G17
G8
C13
C9
G12
G10
C11
chromo domain
G410
N409
R408
I407
R406
L404
β8
β9
B
C
R59
N55
R54
F58
K374
E37
Y370
D
E37
R54
T171
R59
M172
F58
Y370
K374
E
K404
R408
R406
Q65
F
N
α15
L376
T438
T380
W418
β8
M384
α17
C


A
GEN1:HJ
0 8 16 32 64 128 256 (nM)
GEN1:5'Flap
0 0.5 1 2 4 8 16 (nM)
B
HJ Cleavage (%)
5'Flap Cleavage (%)
[GEN1] (nM)
WT
R93E
R89E
H109E
F110E
R54E
D30N
C36E
C
HJ Cleavage (%)
5'Flap Cleavage (%)
[GEN1] (nM)
WT
R406E
T438E
T380E
Δchromo
D
L436
T418
T438
Y424
L397
K404
K406
E440
T380
R54
D30
H109
F110
C36
active site
R93
K219
L436
K404
W418
Y424
T438
R406
L397
T380
E440
R54
H109
C36
F110
D30
R93
active site
E
cruciform
(–)
supercoiled
linear
nicked
F
WT
Δchromo
F110E
C36E
R406E
R54E
R89E
GEN1
nicked
linear
s.c.
G
MMS
YPD
0.0025%
0.005%
0.0075%
0.01%
mus81Δ yen1Δ
YEN1 WT
YEN1 D41N
YEN1 F47E
YEN1 K298E
YEN1 I97E
nuclease
YEN1 K469E
YEN1 F478A
YEN1 K484E
YEN1 R486E
YEN1 Y487A
YEN1 N526A
YEN1 L528D
YEN1 W529A
YEN1 L530A
chromodomain
YEN1 Δchromo


A
hsGEN1+hsFEN1
B
upper gate
lower gate
arch
α4
α14
β8
*
α3
α2
~94°
α6
lower gate
α4
α10
hsGEN1-flap model
C
*
*
hsGEN1-HJ model
D
ttRuvC-HJ
E
T4 endo VII-HJ
F
T7 endo I-HJ