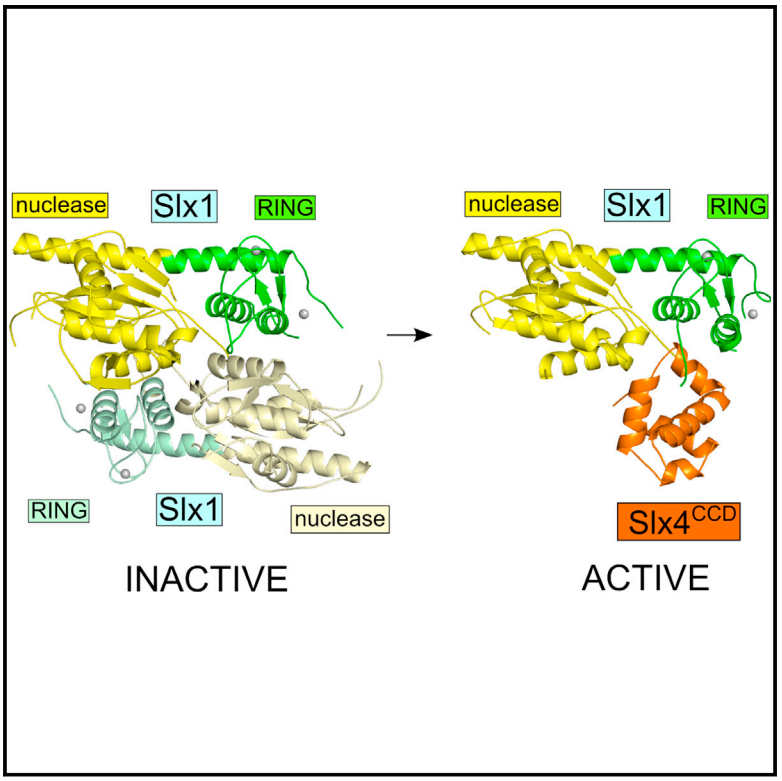
Cell Reports

Report

Structural and Mechanistic Analysis of the Slx1-Slx4 Endonuclease

Graphical Abstract



Authors

Vineet Gaur, Haley D.M. Wyatt, ..., Stephen C. West, Marcin Nowotny

Correspondence

mnowotny@iimcb.gov.pl

In Brief

The SLX1-SLX4 endonuclease cleaves branched DNA structures and has critical roles in DNA replication, recombination, and repair. Gaur et al. report structures for yeast Slx1 alone and in complex with the C-terminal domain of Slx4. These studies reveal a potential regulatory mechanism for the nuclease activity of Slx1.

Highlights

- Structural insights into the Slx1 nuclease and Slx1-Slx4 heterodimer
- Slx1 forms a stable homodimer in which the active site is blocked
- Slx1 homodimerization and interaction with Slx4 are mutually exclusive
- Conversion of Slx1 homodimer to Slx1-Slx4 heterodimer is proposed to activate Slx1

Accession Numbers

4XM5
4XLG

Gaur et al., 2015, Cell Reports *10*, 1467–1476
March 10, 2015
http://dx.doi.org/10.1016/j.celrep.2015.02.019

CellPress

Cell Reports
Report

Structural and Mechanistic Analysis of the Slx1-Slx4 Endonuclease

Vineet Gaur,[1] Haley D.M. Wyatt,[2] Weronika Komorowska,[1] Roman H. Szczepanowski,[3] Daniele de Sanctis,[4] Karolina M. Gorecka,[1] Stephen C. West,[2] and Marcin Nowotny[1,*]
[1]Laboratory of Protein Structure, International Institute of Molecular and Cell Biology, 4 Księcia Trojdena Street, 02-109 Warsaw, Poland
[2]London Research Institute, Cancer Research UK, Clare Hall Laboratories, Blanche Lane, South Mimms, Herts EN6 3LD, UK
[3]Biophysics Core Facility, International Institute of Molecular and Cell Biology, 4 Księcia Trojdena Street, 02-109 Warsaw, Poland
[4]European Synchrotron Radiation Facility (ESRF), 71 Avenue des Martyrs, CS 40220, 38043 Grenoble Cédex 9, France
*Correspondence: mnowotny@iimcb.gov.pl
http://dx.doi.org/10.1016/j.celrep.2015.02.019

SUMMARY

The SLX1-SLX4 endonuclease required for homologous recombination and DNA repair in eukaryotic cells cleaves a variety of branched DNA structures. The nuclease subunit SLX1 is activated by association with a scaffolding protein SLX4. At the present time, little is known about the structure of SLX1-SLX4 or its mechanism of action. Here, we report the structural insights into SLX1-SLX4 by detailing the crystal structure of *Candida glabrata* (*Cg*) Slx1 alone and in combination with the C-terminal region of Slx4. The structure of Slx1 reveals a compact arrangement of the GIY-YIG nuclease and RING domains, which is reinforced by a long α helix. Slx1 forms a stable homodimer that blocks its active site. Slx1-Slx4 interaction is mutually exclusive with Slx1 homodimerization, suggesting a mechanism for Slx1 activation by Slx4.

INTRODUCTION

The human SLX1-SLX4 structure-selective endonuclease plays a key role in DNA repair, homologous recombination, replication fork restart, and telomere maintenance (Svendsen and Harper, 2010). The genes encoding *Saccharomyces cerevisiae* Slx1 and Slx4 were discovered in a genetic screen for mutations that are synthetic lethal in the absence of the Sgs1 helicase, a protein that is important for genome stability (Mullen et al., 2001). Homology searches subsequently identified the *SLX1* and *SLX4* genes in higher eukaryotes (Andersen et al., 2009; Fekairi et al., 2009; Muñoz et al., 2009; Svendsen et al., 2009). Slx1 is an evolutionarily conserved protein that contains an N-terminal GIY-YIG nuclease domain (also called URI domain) and a C-terminal zinc-finger domain. GIY-YIG domains also are present in homing nucleases, the bacterial nucleotide excision-repair nuclease UvrC, and several type II restriction enzymes (Dunin-Horkawicz et al., 2006). The mechanism of substrate binding and cleavage for GIY-YIG family members has been elucidated by crystallographic studies of protein-DNA complexes obtained for two restrictases, namely R.Eco29kI (Mak et al., 2010) and Hpy188I (Sokolowska et al., 2011).

The Slx4 subunit of the Slx1-Slx4 nuclease is thought to provide a scaffold that coordinates the actions of a number of proteins involved in DNA processing (Cybulski and Howlett, 2011). For example, vertebrate SLX4 is a large, multi-domain protein that interacts with several DNA repair proteins (Andersen et al., 2009; Fekairi et al., 2009; Muñoz et al., 2009; Salewsky et al., 2012; Svendsen et al., 2009): (1) the N-terminal region of human SLX4 binds the MSH2-MSH3 mismatch-repair complex and XPF-ERCC1 nucleotide excision-repair enzyme; and (2) the C-terminal portion of SLX4 binds the telomeric proteins TRF2 and RAP1, the PLK1 kinase, and the MUS81-EME1 endonuclease. In all organisms studied to date, SLX1 binds to the extreme C-terminal region of SLX4, which contains an evolutionarily conserved helix-turn-helix motif. Interestingly, in vitro studies have shown that SLX4 stimulates the endonuclease activities of SLX1, MUS81-EME1, and XPF-ERCC1 (Hodskinson et al., 2014; Muñoz et al., 2009; Wyatt et al., 2013). The importance of SLX4 is demonstrated by the observation that biallelic mutations in SLX4 (also known as FANCP) are associated with the cancer-prone disorder Fanconi anemia (Bogliolo et al., 2013; Kim et al., 2011; Stoepker et al., 2011).

Although the amino acid sequence of SLX4 is evolutionarily diverse, the C-terminal region of all SLX4 proteins contains a conserved C-terminal domain (CCD) that underpins the interaction with SLX1 and a DNA-binding SAP domain found in many DNA repair proteins (Andersen et al., 2009; Aravind and Koonin, 2000; Fekairi et al., 2009; Muñoz et al., 2009; Svendsen et al., 2009). In yeast, there are few other discernible domains, whereas SLX4 proteins from higher eukaryotes (e.g., worms, flies, and humans) contain one or two copies of a UBZ family zinc-finger domain known as UBZ4; the MEI9^{XPF} interaction like region (MLR); and a Broad-complex, Tramtrack, and Bric-a-brac (BTB) domain (Stogios et al., 2005).

The role of SLX1-SLX4 in DNA repair has been studied extensively (Cybulski and Howlett, 2011; Sarbajna and West, 2014; Wyatt and West, 2014). Although deletion of Slx1 in yeast does not affect the response to DNA damage, it has been shown that Slx1-Slx4 plays a role in maintaining the integrity of ribosomal loci, which contain tandem repeats that frequently lead to replication fork arrest (Coulon et al., 2004). It is possible that

Slx1-Slx4 is involved in the collapse of stalled forks and the resolution of recombination intermediates, such as Holliday junctions (HJs), after fork recapture (Gritenaite et al., 2014). In human cells, transient depletion of SLX4 leads to an increased sensitivity to alkylating and crosslinking agents, indicating the importance of SLX4 for the repair of DNA inter-strand crosslinks (ICLs) and protein-DNA adducts. Depletion of SLX4 also reduces the efficiency of double-strand break repair and leads to genome instability (Garner et al., 2013; Muñoz et al., 2009; Sarbajna et al., 2014; Svendsen et al., 2009; Wechsler et al., 2011). Collectively, these observations indicate that SLX1 and/or SLX4 have relatively well-conserved roles in processing DNA intermediates that arise at stalled or collapsed replication forks, particularly when cells are treated with DNA-damaging agents that interfere with normal replication fork progression.

Purified *S. cerevisiae* Slx1-Slx4 cleaves various DNA substrates in vitro, including splayed-arm structures, model replication forks, 5′-flaps, and HJs (Fricke and Brill, 2003). For the 5′-flap substrates, the major cleavage site lies in the 5′ single-stranded arm at the junction between single- and double-stranded DNA. Whereas Slx1 alone possesses weak nuclease activity, the rate is stimulated approximately 500-fold by Slx4. Purified human SLX1-SLX4 exhibits related activities and cleaves 5′-flaps, 3′-flaps, replication forks, and HJs. Of particular interest, SLX1-SLX4 and MUS81-EME1 cooperate during HJ resolution, with SLX1-SLX4 performing the initial nick such that MUS81-EME1 can resolve the nicked HJ without substrate dissociation (Wyatt et al., 2013).

Although there is a significant amount of functional information available for Slx1-Slx4 and its associated proteins, structural and mechanistic insights for this enzyme are lacking. Here we report the crystal structure of the Slx1 nuclease, obtained using *Candida glabrata* Slx1 (*Cg*-Slx1). The protein forms a compact structure with the GIY-YIG nuclease and RING-finger domains interacting with each other, and the structural arrangement is reinforced by a long α helix. We find that *Cg*-Slx1 forms a stable homodimer in the absence of *Cg*-Slx4. Importantly, the crystal structure of *Cg*-Slx1 in complex with *Cg*-Slx4^{CCD}, together with biochemical analyses, demonstrate that *Cg*-Slx1 homodimerization is mutually exclusive with the formation of a *Cg*-Slx1-Slx4^{CCD} heterodimer, revealing a likely regulatory mechanism for Slx1 endonuclease activity.

RESULTS AND DISCUSSION

Overall Structure of *Cg*-Slx1

To obtain structural information for Slx1, we purified *C. glabrata* Slx1 protein (*Cg*-Slx1) alone and in complex with the conserved C-terminal domain of Slx4 (*Cg*-Slx1-Slx4^{CCD}). When the enzymatic activities of *Cg*-Slx1 and *Cg*-Slx1-Slx4^{CCD} were verified using synthetic DNA substrates (for sequences, see Figures S1A and S1B), we found that *Cg*-Slx1 exhibited little or no activity, whereas the *Cg*-Slx1-Slx4^{CCD} complex cleaved a diverse set of branched DNA structures (Figure 1). The observed nuclease activity was inherent to Slx1, as substitution of an invariant residue in the active site (Slx1^{E79Q}) abolished the nuclease activity of *Cg*-Slx1-Slx4^{CCD} (Figure 1B, bottom). We observed that the HJ was the preferred substrate for *Cg*-Slx1-Slx4^{CCD}, followed by 5′-flap, splayed arm, and 3′-flap DNA structures (Figures 1B [top] and 1C). In contrast, the enzyme failed to cleave gapped, nicked, and double- or single-stranded DNA substrates. A similar substrate preference was reported for full-length yeast (Fricke and Brill, 2003) and human (Wyatt et al., 2013) enzymes, although *Cg*-Slx1-Slx4^{CCD} was comparatively more active on 3′-flaps. Enzyme titration experiments with the 5′-flap substrate revealed that the primary cleavage product, a nicked or gapped DNA duplex, was processed further to generate a shorter DNA duplex (Figure 1A). However, time-course experiments carried out with limiting amounts of protein revealed that the major reaction was a single nick that converted the 5′-flap into a linear duplex (Figure 1B). This is consistent with the inability of *Cg*-Slx1-Slx4^{CCD} to cleave nicked or gapped DNA duplexes when protein is limiting. To gain more insight into the mechanism by which *Cg*-Slx1-Slx4^{CCD} cleaves 5′-flaps, we prepared two different substrates, each ^{32}P-labeled on a different oligonucleotide (Figures S1C and S1D). Time-course experiments revealed two independent cleavage sites: a primary incision that removed the 5′-flap, and a less efficient incision reaction that removed the duplex arm. Collectively, our results provide additional evidence for the promiscuous nature of Slx1-Slx4, suggesting that this biochemical property has been conserved throughout evolution.

We next obtained crystals of *Cg*-Slx1^{E79Q}, which belonged to the $P4_32_12$ space group, and diffracted X-rays to 2.34 Å resolution. The protein was predicted to contain a zinc finger, so diffraction data were collected at the zinc absorbance peak wavelength (1.280 Å), and the structure was solved by zinc single-wavelength anomalous diffraction (SAD) (Table S1). The asymmetric unit of the crystal contained one Slx1 molecule and the structure was refined to an R_{free} of 25.1% (Table S1). Fragments of simulated annealing composite omit electron density maps are presented in Figure S2A. We noted that the loop regions of the zinc-finger domain had higher B factors and less-well-defined electron density maps, indicating the flexibility of this part of the structure.

The *Cg*-Slx1 monomer consists of two distinct domains: (1) an N-terminal GIY-YIG or URI nuclease domain, and (2) a C-terminal zinc-finger domain. In the *Cg*-Slx1 structure, these two domains interact with each other and form an oblong molecule approximately 70 Å long and 30–40 Å wide (Figure 2). The overall structure is reinforced by a long α helix (helix α6, located between the two domains and comprising residues 176–214) that spans the entire molecule and provides a scaffold for the two domains. The central element of the nuclease domain is a β sheet consisting of five strands arranged β2-β1-β3-β6-β7 (numbered in the amino acid sequence, Figure S2C), with strands β2 and β3 oriented antiparallel to β1, β6, and β7 (Figure 2). The β sheet is flanked by six helices; short helix α1 and long helix α6 are located on one side of the sheet and four helices (α2–α5) are positioned on the other side. The α2–α5 region also contains a β hairpin formed by strands β4 and β5 (Figure 2). Comparisons of *Cg*-Slx1 structure with related GIY-YIG nucleases can be found in the Supplemental Results and in Figures S2D–S2G. The active site of *Cg*-Slx1 is highly conserved with other GIY-YIG nucleases (Figure S2H), and we therefore propose that the catalytic mechanism of Slx1 is identical to that described for Hpy188I (Sokolowska et al., 2011). Details of the organization of the active site and

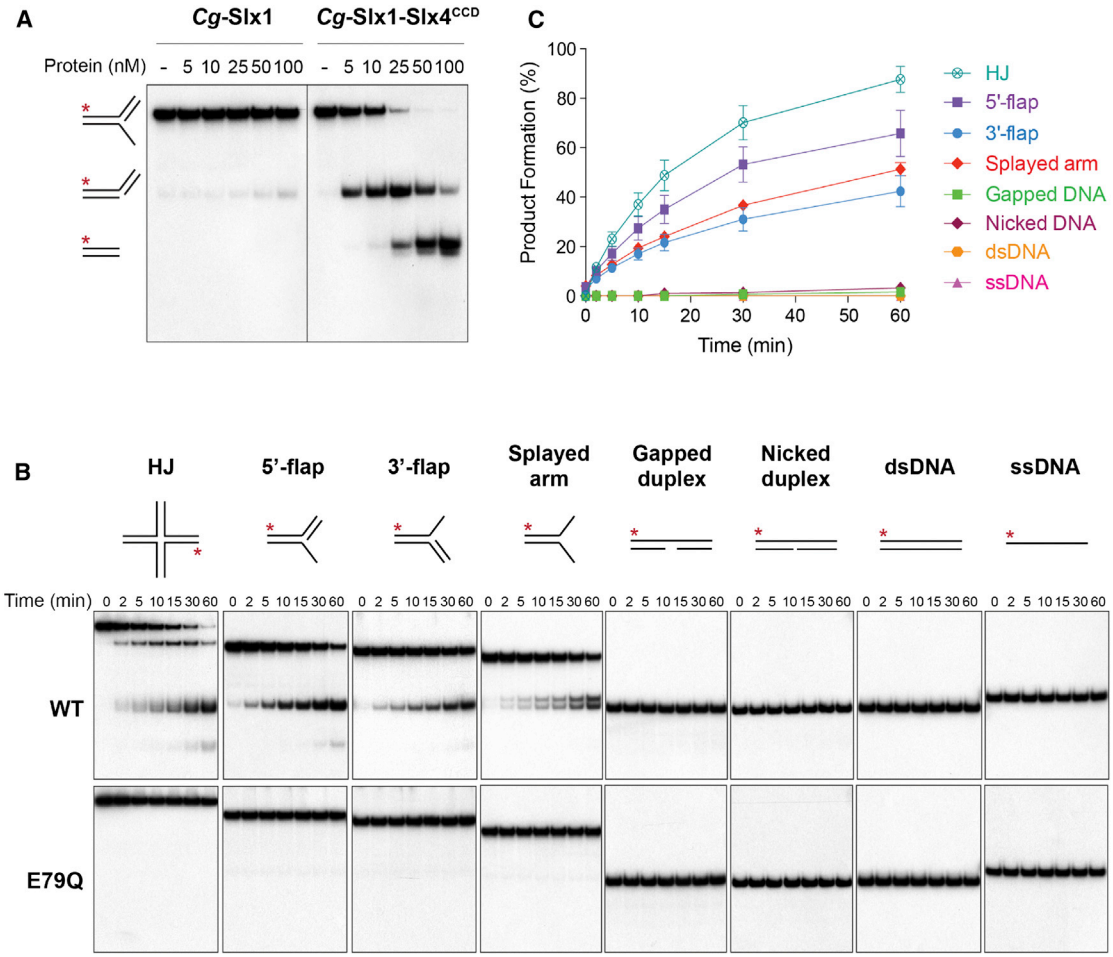


Figure 1. Nuclease Activities of *Cg*-Slx1 and *Cg*-Slx1-Slx4^{CCD}

(A) 5′-flap cleavage by *Cg*-Slx1 and the *Cg*-Slx1-Slx4^{CCD} complex. 5′-flap DNA (100 nM), spiked with negligible amounts of 5′-^{32}P-labeled substrate, was incubated with increasing concentrations of *Cg*-Slx1 or *Cg*-Slx1-Slx4^{CCD} at 37°C for 15 min. Reaction products were analyzed by native PAGE and autoradiography. Red asterisk, 5′-^{32}P-labeled oligonucleotide.

(B) Substrate specificity of *Cg*-Slx1-Slx4^{CCD}. The indicated DNA substrates (100 nM), spiked with trace amounts of the corresponding 5′-^{32}P-labeled substrate, were incubated with 10 nM *Cg*-Slx1-Slx4^{CCD} (top) or 10 nM catalytically inactive *Cg*-Slx1^{E79Q}-Slx4^{CCD} (bottom). At the indicated time points, samples were removed and reaction products were analyzed by native PAGE. Red asterisks, 5′-^{32}P-labeled oligonucleotides.

(C) Quantification of data shown in (B) by phosphorimaging. Product formation is expressed as a percentage of total radiolabeled DNA. Data are presented as the mean of at least three independent measurements ± SEM.

See also Figure S1.

the catalytic mechanism proposed for *Cg*-Slx1 can be found in the Supplemental Results and in Figures S2H and S2I. The C-terminal part of the *Cg*-Slx1 structure comprises a zinc-finger domain. Our structure shows that it is a RING finger closely related to domains found in ubiquitin ligases (Supplemental Results; Figures S2J–S2L).

Slx1 Dimerization

Inspection of the *Cg*-Slx1 crystal packing revealed that a 2-fold crystallographic axis generated a very tight protein dimer, with a total buried surface area of 4,729 Å^2 (Figure 3A). At the dimer interface, helix α2 of one monomer was positioned snugly in a groove located between the nuclease and RING domains of the other subunit. Notably, the side chain of Arg72, found at the N terminus of helix α2, was inserted into a pocket on the surface of the other subunit, where it interacted with the backbone carbonyls of Glu3 and Gln5. Another prominent interaction that stabilized the homodimer involved Tyr86 from the Slx1-specific β4-β5 hairpin of one subunit and helix α7 from the RING domain of the other subunit. Contacts were also made between Glu121, Tyr122, and Tyr86 from one subunit and Thr271, Ile272, and Ile273 of the other. Together, these residues formed a network of charged and van der Waals interactions through their side chains. Two symmetrical copies of this part of the interface were located close to each other and formed a continuous zipper of interdigitated amino acid side chains (Figure 3B).

The extensive dimerization contacts and large buried surface area prompted us to investigate whether *Cg*-Slx1 dimerization

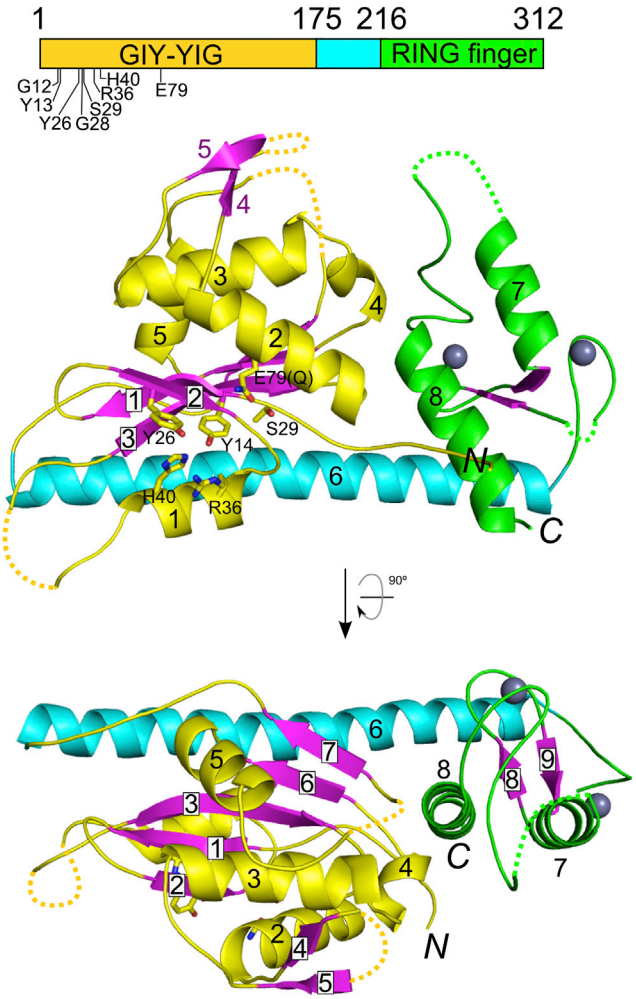


Figure 2. Overall Structure of *Cg*-Slx1

The GIY-YIG nuclease domain is shown in purple for β strands and yellow for the rest of the sequence. The RING domain is shown in green and the connecting scaffold helix α6 in cyan. Zinc ions are shown as gray spheres. The residues forming the active site are shown as sticks and labeled. Dotted lines indicate the loop regions, which were not observed in the electron density maps. Domain boundaries are schematically shown as rectangles on top of the panels with the residue numbers given and active site residues indicated.
See also Table S1 and Figure S2.

occurs in solution. To do this, we used gel filtration coupled to multi-angle light scattering (GF-MALS) (Figure 3C) and analytical ultracentrifugation (AUC) (Figure S3A). The molecular weight (MW) of *Cg*-Slx1 measured by GF-MALS was 80.3 kDa and 72.0 kDa by AUC, suggesting that *Cg*-Slx1 exists as a homodimer in solution (calculated MW of 72.0 kDa). We subsequently investigated the multimeric state of the *Cg*-Slx1-Slx4^{CCD} complex. First, we used GF-MALS to determine that the MW of *Cg*-Slx4^{CCD} alone was 19.0 kDa. This is in agreement with the theoretical MW of 19.7 kDa and indicates that *Cg*-Slx4^{CCD} is monomeric in solution. For *Cg*-Slx1-Slx4^{CCD}, the measured MW was 57.0 kDa (GF-MALS) and 47.0 kDa (AUC). Importantly, these data indicate that the *Cg*-Slx1-Slx4^{CCD} complex contains one molecule of Slx1 and one molecule of Slx4^{CCD} (calculated MW of 57.0 kDa). Our results therefore demonstrate that *Cg*-Slx1 alone exists as a stable homodimer, but, in the presence of *Cg*-Slx4^{CCD}, interacts with this subunit to form a stable heterodimeric complex. Close inspection of the *Cg*-Slx1 homodimeric structure showed that the active site of each subunit was partially blocked by the RING domain of the other subunit (Figures 3A and S4). This suggests that dimerization blocks the activity of *Cg*-Slx1. Indeed, *Cg*-Slx1 (Figure 1A) as well as the *S. cerevisiae* (Fricke and Brill, 2003) and human (Muñoz et al., 2009) enzymes are inactive in the absence of Slx4.

To gain further insight into the *Cg*-Slx1 homodimerization, we prepared a series of *Cg*-Slx1 variants with substitutions in dimer interface residues. These protein variants were structurally unstable (Supplemental Results; Figures S3A–S3C). Additional gel filtration experiments demonstrated that *Cg*-Slx1 homodimerization and *Cg*-Slx1-Slx4^{CCD} complex formation are mutually exclusive, and the exchange between the two forms can be promoted by high salt concentration (Supplemental Results; Figures S3D and S3E).

In summary, our results show that *Cg*-Slx1 is a homodimer in solution. Furthermore, they indicate that *Cg*-Slx1 dimerization provides a physical block to the active site. Thus, these findings provide a potential mechanism by which Slx1 nuclease activity is inhibited.

A Model for DNA Substrate Binding

The *Cg*-Slx1 structure presented here does not contain a DNA substrate. Nonetheless, *Cg*-Slx1-DNA interactions can be modeled based on the conservation of the active site between Slx1 and GIY-YIG restrictases, for which high-resolution protein-DNA structures are available. To this end, we superimposed the Eco29kI structures (Protein Data Bank [PDB] ID 3NIC [Mak et al., 2010]) on our *Cg*-Slx1 structure, which resulted in a good fit between DNA from the Eco29kI structures and the surface of Slx1. The modeling revealed that *Cg*-Slx1 contains two groups of potential DNA-binding residues, all but one of which are located in the nuclease domain (Figure 4A). The first group (Arg35, Arg38, Gln39, and Gln191) could interact with the non-cleaved DNA strand toward its 5′ end from the active site. Based on the polarity of the DNA at the active site, these residues are predicted to contact regions of flap and splayed arm substrates that contain only double-stranded DNA (Figure 4B). Arg35, Arg38, and Gln39 were all located in the vicinity of the phosphate groups of the modeled DNA. Sequence alignment of fungal Slx1 proteins revealed that Arg35 and Gln39 are strictly conserved and Arg38 is partially conserved (Figure S2C). Non-conserved Gln191 was also located in this region but was positioned farther from the substrate in our *Cg*-Slx1-DNA model.

The second group of predicted DNA-binding residues (Arg72, Gln77, His80, and His84) was located on the opposite face of the active site relative to the first group. These residues are predicted to contact the 5′ region of the cleaved strand that exists as single-stranded DNA in flap and splayed arm substrates. In the Eco29kI-DNA structure, the straight DNA duplex extended farther away from these residues, suggesting that the single-stranded portion of the splayed arm or 5′-flap substrates must be bent to interact with the protein. Gln77, His80, and His84

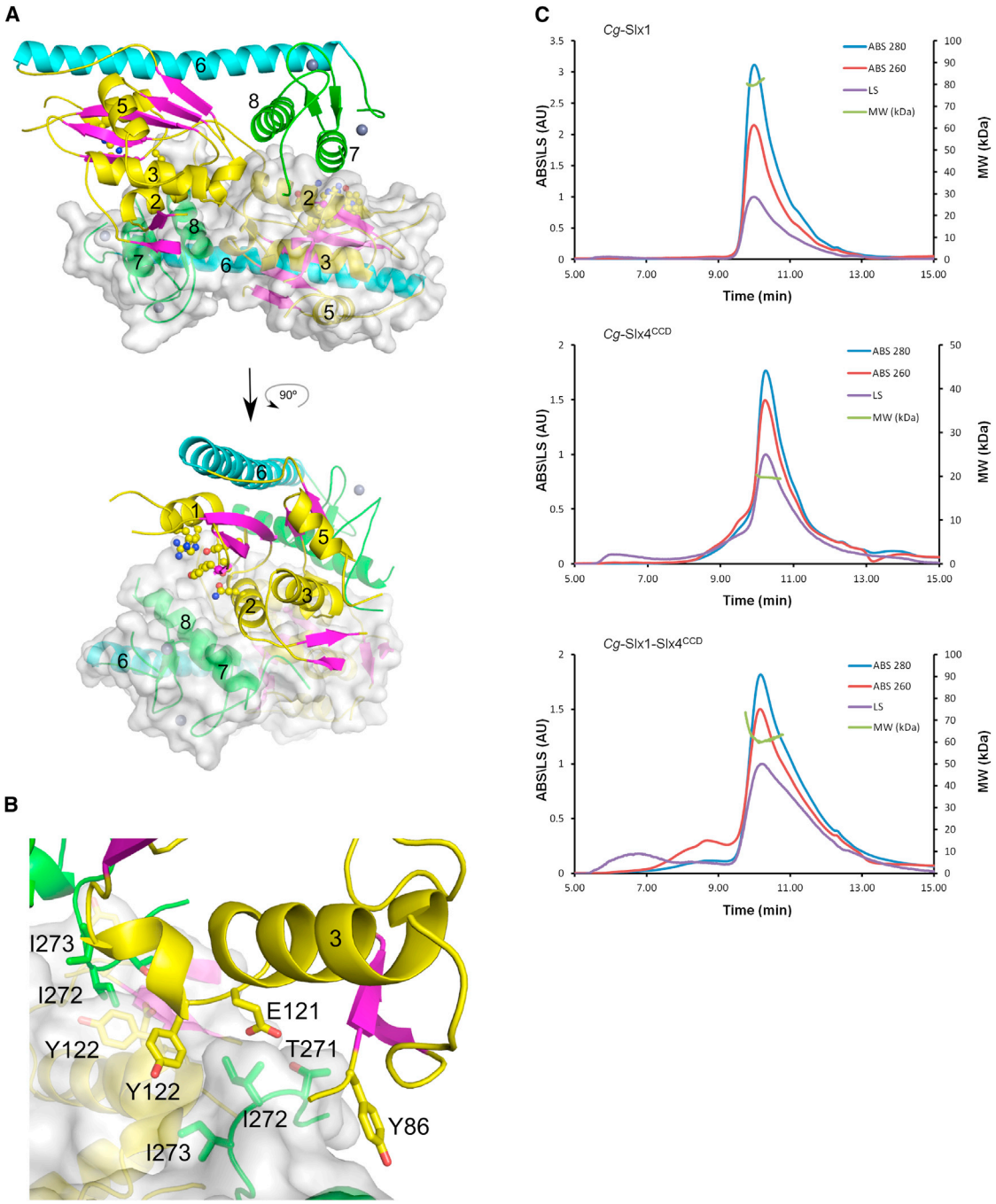


Figure 3. Slx1 Dimerization

(A) One molecule is shown in cartoon representation and colored as in Figure 2. The other subunit is shown in cartoon and transparent surface. The residues forming the active site are shown as ball and sticks.

(B) Close-up of the central region of the dimer interfaces.

(C) Estimation of the MW of *Cg*-Slx1, *Cg*-Slx4^{CCD}, and *Cg*-Slx1-Slx4^{CCD} using GF-MALS. The elution traces from the silica gel filtration column are shown: absorbance at 260 nm (ABS 260) as red trace, at 280 nm (ABS 280) as blue trace, and light scattering (LS) as violet trace. The measured MW throughout the peak is plotted in green (MW values are on the right axis). See also Figure S3.

were all located in helix α2 very close to the metal-coordinating active site residue Glu79. Among fungal Slx1 proteins, Gln77 is conserved and, in some species, replaced with a lysine. Notably, His80 and His84 are less conserved. Although His80 is substituted by a tryptophan in most species, His84 is always replaced with a charged residue. Interestingly, both Arg72 and the

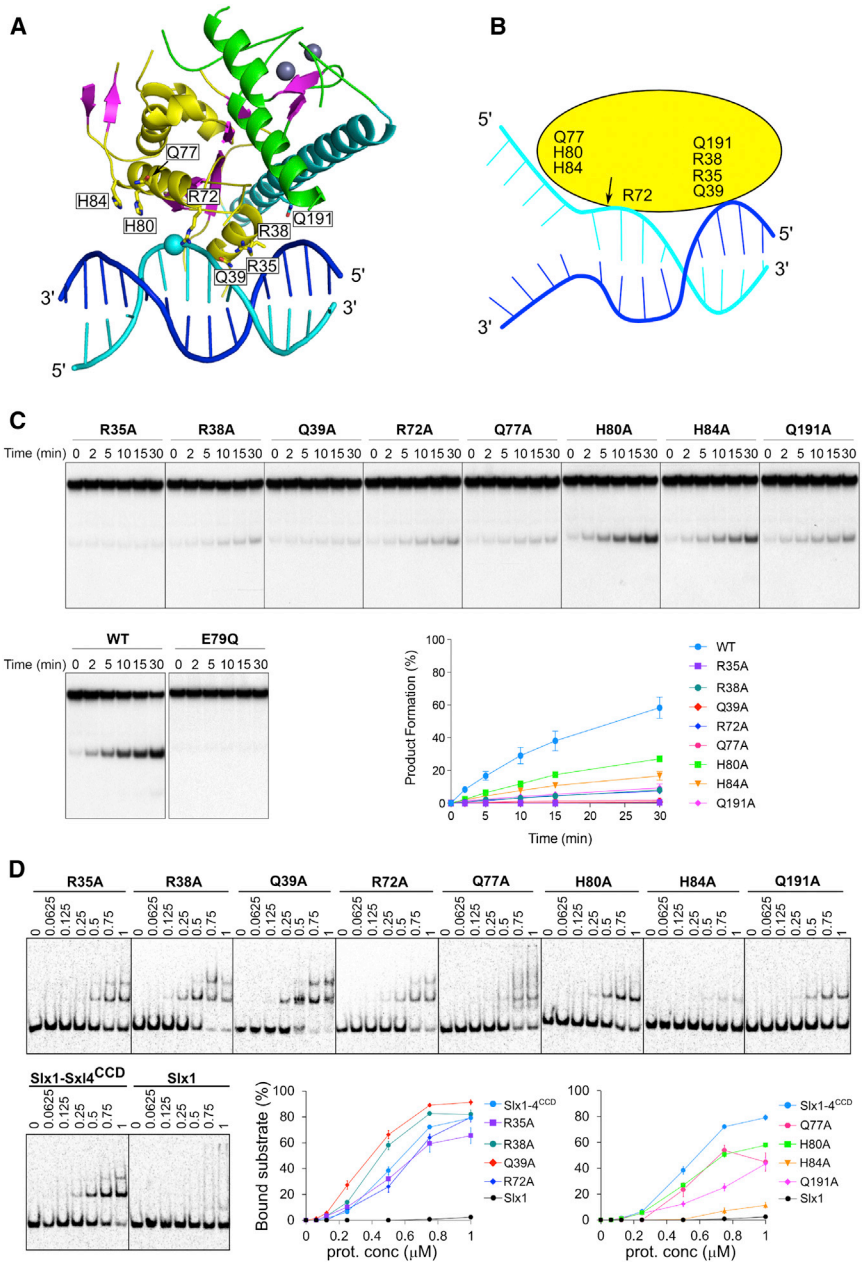


Figure 4. Model of *Cg*-Slx1-DNA Interactions

(A) A view presenting modeled DNA (cyan for the cleaved strand and blue for the non-cleaved one) interacting with *Cg*-Slx1. DNA was modeled by superimposing the catalytic cores of *Cg*-Slx1 and Eco29kI (PDB ID 3NIC [Mak et al., 2010]). The color scheme of *Cg*-Slx1 is the same as in Figure 2. Amino acids predicted to interact with the DNA substrate are shown as sticks. The sphere indicates the scissile phosphate.

(B) Cartoon of the predicted protein-DNA interactions. The site of nucleolytic cleavage is indicated with an arrow.

(C) Cleavage of 5′-flap substrates by *Cg*-Slx1-Slx4^{CCD} mutants that contain single alanine substitutions of residues predicted to interact with DNA. The wild-type *Cg*-Slx1-Slx4^{CCD} and nuclease-inactive *Cg*-Slx1^{E79Q}-Slx4^{CCD} complexes were used as positive and negative controls, respectively. 5′-flap DNA (100 nM) was spiked with negligible amounts of 5′-^{32}P-labeled substrate and incubated with the indicated *Cg*-Slx1-Slx4^{CCD} complex (10 nM) at 37°C for 0, 2, 5, 10, 15, or 30 min. Reaction products were analyzed by native PAGE. Cleavage products were quantified by phosphorimaging and plotted as the percentage of total radiolabelled DNA. Data are presented as the mean of three independent experiments ± SEM.

(D) DNA binding of *Cg*-Slx1 and *Cg*-Slx1-Slx4^{CCD} (wild-type and mutants) determined by EMSA. Protein, as indicated on the top (μM), was added to 5′-flap substrate (20 nM 5′-^{33}P-labeled DNA and 105 nM cold DNA). The mixtures were resolved on 6% native TBE gels and visualization by autoradiography. The plots show the percentage of bound substrate. Data are presented as the mean of three independent experiments ± SEM. See also Figure S4.

region around His80 are buried in the *Cg*-Slx1 homodimer, which likely contributes to its inhibition (Figure S4A).

To verify the importance of the potential DNA-binding residues identified in our model, we prepared several *Cg*-Slx1 variants in which these residues were individually substituted with alanine. The *Cg*-Slx1 mutants were purified in the context of the *Cg*-Slx1-Slx4^{CCD} complex and enzyme activity was determined using the 5′-flap substrate (Figure 4C). These mutants showed either a complete loss of nuclease activity (R35A and Q39A) or substantially impaired activity (R38A, R72A, Q77A, Q191A, H80, and H84). We also tested the ability of these Slx1 variants to bind the 5′-flap substrate using electrophoretic mobility shift assays (EMSAs) (Figure 4D). *Cg*-Slx1 alone did not exhibit DNA-binding properties, while *Cg*-Slx1-Slx4^{CCD} formed a protein-DNA complex with reduced mobility. At high protein concentrations, two protein-DNA complexes were observed, possibly indicating the binding of two Slx1 molecules to DNA. The R38A and Q39A mutants showed slightly stronger substrate binding compared to the wild-type protein, despite exhibiting severe catalytic defects. Importantly, Arg38 and Gln39 are located close to the active site and Gln39 forms a hydrogen bond with active site residue Arg36. Therefore, Arg38 and Gln39 likely participate in active site stabilization upon DNA binding. The R72A variant displayed nearly wild-type binding affinity for the substrate. Its greatly reduced activity suggests that Arg72, which is located in the vicinity of the scissile phosphate, has a more important role in aligning the substrate for cleavage than enhancing substrate affinity. The remaining mutants, namely R35A, Q77A, H80A, H84A, and Q191A, all displayed defects in DNA binding. Interestingly, the most severe defects were observed with the H84A variant.

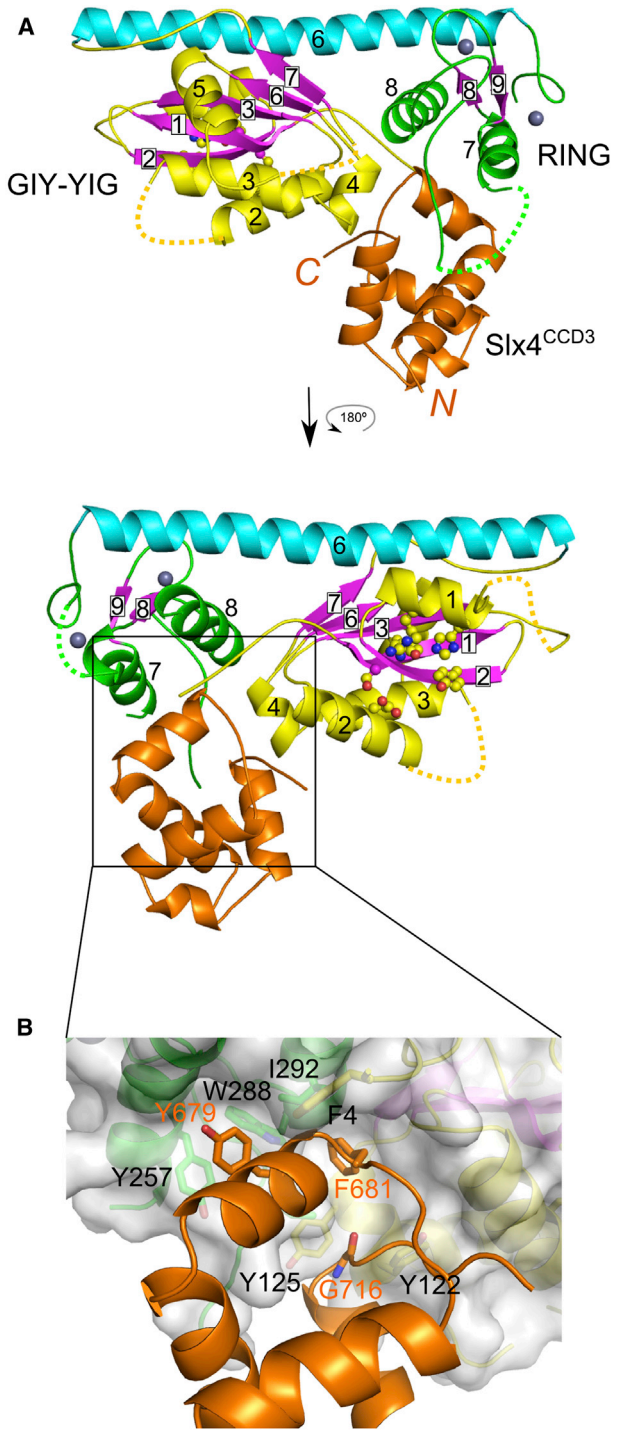


Figure 5. Structure of *Cg*-Slx1-Slx4^{CCD3} Complex

(A) Two views of the complex structure. *Cg*-Slx1 is colored as in Figure 2 and *Cg*-Slx4^{CCD3} is shown in orange. The residues forming the active site are shown as ball and sticks. Dotted lines indicate the loop regions, which were not observed in the electron density maps.

(B) Close-up view of the interactions that occur within the *Cg*-Slx1-Slx4^{CCD3} complex. Selected residues from the interface, including Gly716 from the C terminus of the last helix of CCD3, are shown as sticks.

See also Figure S5.

His84 is located in helix α2, which we postulate to participate in binding the single-stranded portion of the 5′-flap DNA. This result further underscores the importance of helix α2 in the DNA-binding interface.

Collectively, the biochemical results support our proposed model of Slx1-DNA interactions. We have therefore identified two regions of *Cg*-Slx1 that are important for protein-DNA interactions: (1) one predicted to bind double-stranded portion of the DNA, and (2) the other potentially interacting with the single-stranded flap.

Slx1-Slx4^{CCD3} Complex Structure

Our initial crystallization trials with *Cg*-Slx1-Slx4^{CCD} did not yield any crystals, so to gain further insight into *Cg*-Slx1-Slx4 interaction, we performed additional deletion studies. Based on bioinformatic predictions and limited proteolysis experiments, we designed four truncated forms of Slx4^{CCD}, comprising residues 557–685 (*Cg*-Slx4^{CCD1}), 608–685 (*Cg*-Slx4^{CCD1A}), 557–698 (*Cg*-Slx4^{CCD2}), or 647–726 (*Cg*-Slx4^{CCD3}) (Figure S5A). All of these fragments were soluble when expressed in *Escherichia coli* (Figure S5B), but only the CCD3 variant co-purified with *Cg*-Slx1 (Figures S5C and S5D). When we analyzed the activity of *Cg*-Slx1-Slx4^{CCD3} on the 5′-flap DNA, we found that it displayed slightly less activity than *Cg*-Slx1-Slx4^{CCD} and that the cleavage sites for both CCD variants were similar (Figure S5E).

We subsequently obtained crystals of the *Cg*-Slx1-Slx4^{CCD3} complex. They belonged to the $P6_3$ space group and diffracted X-rays to 1.8 Å (Table S1). The structure was solved by molecular replacement using the *Cg*-Slx1 model and the structure of *Cg*-Slx4^{CCD3} was traced manually (Figures 5 and S2B). The overall conformation of *Cg*-Slx1 domains did not change significantly upon Slx4 binding, although the location of the RING domain changed slightly (Figures S5F and S5G). When *Cg*-Slx1 and *Cg*-Slx1-Slx4^{CCD3} were superimposed using the GIY-YIG domains, the position of the RING domain residues differed by up to ~3.5 Å. The β4-β5 hairpin was not formed in the *Cg*-Slx1-Slx4^{CCD3} complex and this part of the structure adopted a different conformation (Figure 5A). This suggests that the hairpin was stabilized by *Cg*-Slx1 homodimerization.

The structure of *Cg*-Slx4^{CCD3} comprises five α helices. Its closest structural homologs are FF domains mediating protein-protein interactions (Bedford and Leder, 1999). For example, Prp40 FF1 domain (PDB ID 2B7E [Gasch et al., 2006]) can be superimposed on *Cg*-Slx4^{CCD3} with an rmsd of 1.6 Å over 37 Cα atoms. Notably, however, the characteristic phenylalanine residues of the FF domain are not conserved in *Cg*-Slx4. *Cg*-Slx4^{CCD3} was positioned in a cleft between the nuclease and RING domains, and most of the interactions between the *Cg*-Slx1 and *Cg*-Slx4^{CCD3} occurred through hydrophobic contacts (Figure 5B). In particular, the second short α helix of *Cg*-Slx4^{CCD3} was placed snugly in a hydrophobic groove on the *Cg*-Slx1 surface. The groove lined up with the side chains of *Cg*-Slx1 Phe4 (nuclease domain) and Tyr257, Ile292, and Ile273 (RING domain). The side chain of *Cg*-Slx4 Phe681, which is conserved in most fungal Slx4 sequences (Figure S5A) and is located in the second short α helix, was inserted into a pocket on the *Cg*-Slx1 surface. The hydrophobic interactions were reinforced further by stacking of *Cg*-Slx1 Tyr257 with

Cg-Slx4 Tyr679. In addition, the backbone carbonyl of this tyrosine formed a hydrogen bond with the Nε1 of the conserved Trp288 in *Cg*-Slx1. Additional residues at the *Cg*-Slx1-Slx4 interface included *Cg*-Slx1 Ile73, Tyr122, and Tyr125, all of which interacted with the C terminus of the last α helix of *Cg*-Slx4^{CCD3} and the following loop (residues 714–720). The importance of the latter interaction was confirmed by the lack of binding between *Cg*-Slx1 and Slx4^{CCD2}, in which the C-terminal helix of the CCD is absent.

In murine SLX4, the C1536R mutation abolished its interaction with SLX1 (Castor et al., 2013). This cysteine corresponds to *Cg*-Slx4 Gly716 located at the C terminus of the last helix of *Cg*-Slx4^{CCD3}. Insertion in this position of an arginine residue with a large side chain would lead to steric clashes with *Cg*-Slx1, explaining the effect of C1536R mutation. This observation further validates the interactions identified in our *Cg*-Slx1-Slx4^{CCD3} structure. Importantly, the position of *Cg*-Slx4^{CCD3} in the Slx1-Slx4 heterodimer overlapped with the position of one of the Slx1 monomers in the *Cg*-Slx1 homodimer (Figures S5F and S5G). In fact, several residues, including *Cg*-Slx1 Tyr122 and Ile273, participated in both *Cg*-Slx1 homodimerization and *Cg*-Slx1-Slx4 heterodimerization, thus explaining why the formation of *Cg*-Slx1 homodimeric and *Cg*-Slx1-Slx4^{CCD} heterodimeric complexes are mutually exclusive events. Slx4^{CCD3} was located away from the predicted DNA-binding interface of *Cg*-Slx1 and is not expected to form contacts with the substrate (Figure S4B).

In conclusion, the studies presented here provide the first structural information for Slx1 and its interaction with Slx4. Structure-selective endonucleases pose an inherent threat to genome integrity because broken DNA ends can facilitate chromosome rearrangements and genome alterations that are potentially tumorigenic. It is therefore crucial to keep these endonucleases tightly regulated to ensure cleavage of the correct DNA substrate at an appropriate time in the cell cycle. Recent studies have revealed that these regulatory mechanisms operate at various levels: protein expression (Courcelle et al., 2001), post-translation modification (Mus81-Eme1/Mms4, Yen1) (Blanco et al., 2014; Matos et al., 2011, 2013), nuclear localization (GEN1) (Chan and West, 2014), conformational changes (FEN1) (Kim et al., 2001; Storici et al., 2002), and protein-protein interactions (e.g., XPG, XPF-ERCC1, and SLX-MUS) (Araújo et al., 2001; Li et al., 1994; Wyatt et al., 2013; Zotter et al., 2006). Our data suggest a mechanism of Slx1 regulation through inhibitory homodimerization that would keep Slx1, a promiscuous and potentially dangerous endonuclease, latent and ensure that its activity is tightly regulated in cells.

EXPERIMENTAL PROCEDURES

Crystallization

Purified *Cg*-Slx1, at a concentration of 10 mg/ml, was subjected to crystallization screens at 18°C using the sitting-drop vapor diffusion method. Prior to crystallization, the protein was dialyzed against buffer containing 20 mM HEPES-NaOH (pH 7.5), 350 mM NaCl, and 1 mM DTT. Crystals were obtained with 1.4 M sodium citrate tribasic dihydrate and 0.1 M HEPES-NaOH (pH 7.5). Crystals were cryoprotected using 30% glycerol (v/v) before data collection. Purified *Cg*-Slx1-Slx4^{CCD3}, at a concentration of 10 mg/ml, was subjected to crystallization screens at room temperature using the sitting-drop vapor diffusion method. Prior to crystallization, the protein was dialyzed against buffer containing 20 mM HEPES-NaOH (pH 7.5), 150 mM NaCl, and 1 mM DTT. Crystals were obtained with 0.2 M magnesium chloride hexahydrate, 0.1 M Bis Tris (pH 6.5), and 25% (v/v) PEG 3350 in the presence of a splayed arm DNA substrate with 15-bp double-stranded portion and 5-nt single-stranded overhangs. Crystals were cryoprotected using 30% glycerol (v/v) before data collection.

Diffraction Data Collection, Structure Solution, and Refinement

X-ray diffraction data for *Cg*-Slx1 were collected at beamline 14.1 at Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung (BESSY) (Mueller et al., 2012), and data for *Cg*-Slx1-Slx4^{CCD3} were collected at beamline ID29 at European Synchrotron Radiation Facility (ESRF). Data for the phasing and refinement of *Cg*-Slx1 structure were collected at 1.280 Å wavelength. Diffraction data used for refinement of the *Cg*-Slx1-Slx4^{CCD3} structure were collected at 0.97626 Å. Diffraction data were processed and scaled with XDS (Kabsch, 2010). The statistics of the diffraction data are summarized in Table S1. Phases for *Cg*-Slx1 were determined using single-wavelength anomalous diffraction in AutoSol module in Phenix (Adams et al., 2010), using data collected at Zn peak wavelength (1.280 Å). Phases for *Cg*-Slx1-Slx4^{CCD3} were determined by molecular replacement using Phaser-MR module in Phenix (Adams et al., 2010). The structure of *Cg*-Slx1 was used as the starting model for molecular replacement. Interactive model building was performed in COOT (Emsley et al., 2010) and refinement with Phenix with R-free, calculated with 5% of unique reflections. The structures of *Cg*-Slx1 and *Cg*-Slx1-Slx4^{CCD3} were refined with 95.7% and 99.7% residues in the favored region of the Ramachandran plot, respectively. Structure validation was carried out using Molprobity analysis (Chen et al., 2010). Structural analyses, including superpositions, and structural figures were prepared in Pymol (http://www.pymol.org). Simulated annealing composite omit maps were calculated in CNS 1.3 (Brünger et al., 1998).

Description of protein purification, oligomeric state analysis, nuclease assays, and EMSA can be found in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The Protein Data Bank accession numbers for the structures of *Cg*-Slx1 and *Cg*-Slx1-Slx4^{CCD3} complex reported in this paper are 4XM5 and 4XLG, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Results, Supplemental Experimental Procedures, five figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.02.019.

AUTHOR CONTRIBUTIONS

V.G. and W.K. purified and crystallized the proteins. H.D.M.W. performed nuclease assays. V.G. and D.d.S. collected diffraction data. V.G. and M.N. solved and analyzed the structures. V.G., W.K., and R.H.S. performed biophysical characterization. V.G., W.K., and K.M.G. performed initial cloning and protein production tests. M.N. and S.C.W. supervised the project. V.G., H.D.M.W., S.C.W., and M.N. wrote the manuscript.

ACKNOWLEDGMENTS

We thank Professor Matthias Bochtler for critical reading of the manuscript and I. Ptasiewicz for excellent technical assistance. We would like to thank the staff of beamline 14-1 at Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung (BESSY) for assistance with data collection. This work was supported by a Wellcome Trust International Senior Research Fellowship to M.N. (098022) and by Cancer Research UK, the European Research Council, and the Jeantet Foundation (to S.C.W.). M.N. is a recipient of the Foundation for Polish Science Ideas for Poland award. The research of M.N. was supported in part by an International Early Career Scientist grant from the Howard Hughes Medical Institute. The research was performed using Centre for

Preclinical Research and Technology (CePT) infrastructure (European Union POIG.02.02.00-14-024/08-00 project).

Received: October 28, 2014
Revised: January 16, 2015
Accepted: February 3, 2015
Published: March 5, 2015

REFERENCES

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

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

Araújo, S.J., Nigg, E.A., and Wood, R.D. (2001). Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome. Mol. Cell. Biol. *21*, 2281–2291.

Aravind, L., and Koonin, E.V. (2000). SAP - a putative DNA-binding motif involved in chromosomal organization. Trends Biochem. Sci. *25*, 112–114.

Bedford, M.T., and Leder, P. (1999). The FF domain: a novel motif that often accompanies WW domains. Trends Biochem. Sci. *24*, 264–265.

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

Bogliolo, M., Schuster, B., Stoepker, C., Derkunt, B., Su, Y., Raams, A., Trujillo, J.P., Minguillón, J., Ramírez, M.J., Pujol, R., et al. (2013). Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. Am. J. Hum. Genet. *92*, 800–806.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

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

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

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Coulon, S., Gaillard, P.H., Chahwan, C., McDonald, W.H., Yates, J.R., 3rd, and Russell, P. (2004). Slx1-Slx4 are subunits of a structure-specific endonuclease that maintains ribosomal DNA in fission yeast. Mol. Biol. Cell *15*, 71–80.

Courcelle, J., Khodursky, A., Peter, B., Brown, P.O., and Hanawalt, P.C. (2001). Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient Escherichia coli. Genetics *158*, 41–64.

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

Dunin-Horkawicz, S., Feder, M., and Bujnicki, J.M. (2006). Phylogenomic analysis of the GIY-YIG nuclease superfamily. BMC Genomics *7*, 98.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

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

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

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

Gasch, A., Wiesner, S., Martin-Malpartida, P., Ramirez-Espain, X., Ruiz, L., and Macias, M.J. (2006). The structure of Prp40 FF1 domain and its interaction with the crn-TPR1 motif of Clf1 gives a new insight into the binding mode of FF domains. J. Biol. Chem. *281*, 356–364.

Gritenaite, D., Princz, L.N., Szakal, B., Bantele, S.C., Wendeler, L., Schilbach, S., Habermann, B.H., Matos, J., Lisby, M., Branzei, D., and Pfander, B. (2014). A cell cycle-regulated Slx4-Dpb11 complex promotes the resolution of DNA repair intermediates linked to stalled replication. Genes Dev. *28*, 1604–1619.

Hodskinson, M.R., Silhan, J., Crossan, G.P., Garaycoechea, J.I., Mukherjee, S., Johnson, C.M., Schärer, O.D., and Patel, K.J. (2014). Mouse SLX4 is a tumor suppressor that stimulates the activity of the nuclease XPF-ERCC1 in DNA crosslink repair. Mol. Cell *54*, 472–484.

Kabsch, W. (2010). Xds. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kim, C.Y., Park, M.S., and Dyer, R.B. (2001). Human flap endonuclease-1: conformational change upon binding to the flap DNA substrate and location of the Mg2+ binding site. Biochemistry *40*, 3208–3214.

Kim, Y., Lach, F.P., Desetty, R., Hanenberg, H., Auerbach, A.D., and Smogorzewska, A. (2011). Mutations of the SLX4 gene in Fanconi anemia. Nat. Genet. *43*, 142–146.

Li, L., Elledge, S.J., Peterson, C.A., Bales, E.S., and Legerski, R.J. (1994). Specific association between the human DNA repair proteins XPA and ERCC1. Proc. Natl. Acad. Sci. USA *91*, 5012–5016.

Mak, A.N., Lambert, A.R., and Stoddard, B.L. (2010). Folding, DNA recognition, and function of GIY-YIG endonucleases: crystal structures of R.Eco29kI. Structure *18*, 1321–1331.

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

Matos, J., Blanco, M.G., and West, S.C. (2013). Cell-cycle kinases coordinate the resolution of recombination intermediates with chromosome segregation. Cell Rep. *4*, 76–86.

Mueller, U., Darowski, N., Fuchs, M.R., Förster, R., Hellmig, M., Paithankar, K.S., Pühringer, S., Steffien, M., Zocher, G., and Weiss, M.S. (2012). Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. J. Synchrotron Radiat. *19*, 442–449.

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

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

Salewsky, B., Schmiester, M., Schindler, D., Digweed, M., and Demuth, I. (2012). The nuclease hSNM1B/Apollo is linked to the Fanconi anemia pathway via its interaction with FANCP/SLX4. Hum. Mol. Genet. *21*, 4948–4956.

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

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

Sokolowska, M., Czapinska, H., and Bochtler, M. (2011). Hpy188I-DNA pre- and post-cleavage complexes—snapshots of the GIY-YIG nuclease mediated catalysis. Nucleic Acids Res. *39*, 1554–1564.

Stoepker, C., Hain, K., Schuster, B., Hilhorst-Hofstee, Y., Rooimans, M.A., Steltenpool, J., Oostra, A.B., Eirich, K., Korthof, E.T., Nieuwint, A.W., et al.

(2011). SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. Nat. Genet. *43*, 138–141.

Stogios, P.J., Downs, G.S., Jauhal, J.J., Nandra, S.K., and Privé, G.G. (2005). Sequence and structural analysis of BTB domain proteins. Genome Biol. *6*, R82.

Storici, F., Henneke, G., Ferrari, E., Gordenin, D.A., Hübscher, U., and Resnick, M.A. (2002). The flexible loop of human FEN1 endonuclease is required for flap cleavage during DNA replication and repair. EMBO J. *21*, 5930–5942.

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

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

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

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

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

Zotter, A., Luijsterburg, M.S., Warmerdam, D.O., Ibrahim, S., Nigg, A., van Cappellen, W.A., Hoeijmakers, J.H., van Driel, R., Vermeulen, W., and Houtsmuller, A.B. (2006). Recruitment of the nucleotide excision repair endonuclease XPG to sites of UV-induced dna damage depends on functional TFIIH. Mol. Cell. Biol. *26*, 8868–8879.

Cell Reports

Supplemental Information

Structural and Mechanistic Analysis of the Slx1-Slx4 Endonuclease

Vineet Gaur, Haley D.M. Wyatt, Weronika Komorowska, Roman H. Szczepanowski, Daniele de Sanctis, Karolina M. Gorecka, Stephen C. West, and Marcin Nowotny

SUPPLEMENTAL TABLE AND FIGURES

Table S1. Data collection and refinement statistics. Related to Figure 2.

	Cg-Slx1	*Cg*-Slx1-Slx4^{CCD3}
Data collection		
Space group	$P4_32_12$	$P6_3$
Cell dimensions		
a, *b*, *c* (Å)	57.5, 57.5, 183.6	135.8, 135.8, 56.7
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 120.0
Resolution (Å)	50.00-2.34 (2.48 – 2.34)[a]	40.00-1.78 (1.84 – 1.78)[a]
R_{merge}[b]	9.8 (67.1)	8.1 (124.8)
$I / \sigma I$	13.3 (2.4)	16.3 (2.0)
Completeness (%)	99.9 (99.7)	100.0 (100.0)
Redundancy	6.9 (7.0)	10.2 (10.4)
Refinement		
Resolution (Å)	50.00-2.34	40.00-1.78
No. reflections	24808 (2495)	57362 (2818)
R_{work} / R_{free}[c]	0.1912/0.2506	0.1578/0.1818
No. atoms	2207	3006
Protein	2101	2700
Ligand/ion	3	4
Water	103	302
B-factors	57.9	38.8
Protein	58.0	37.7
Ligand/ion	105.7	37.4
Water	55.1	48.9
R.m.s. deviations		
Bond lengths (Å)	0.009	0.017
Bond angles (°)	1.11	1.46

[a] Values in parentheses are for highest-resolution shell.

[b] $R_{merge} = \Sigma|I - \langle I \rangle|/\Sigma I$, where I is observed intensity and $\langle I \rangle$ is average intensity obtained from multiple observations of symmetry related reflections.

[c] $R_{work} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are observed and calculated structure factor amplitudes respectively. $R_{free} = R_{work}$, calculated using random reflections omitted from the refinement (number of reflections given in parentheses).

A

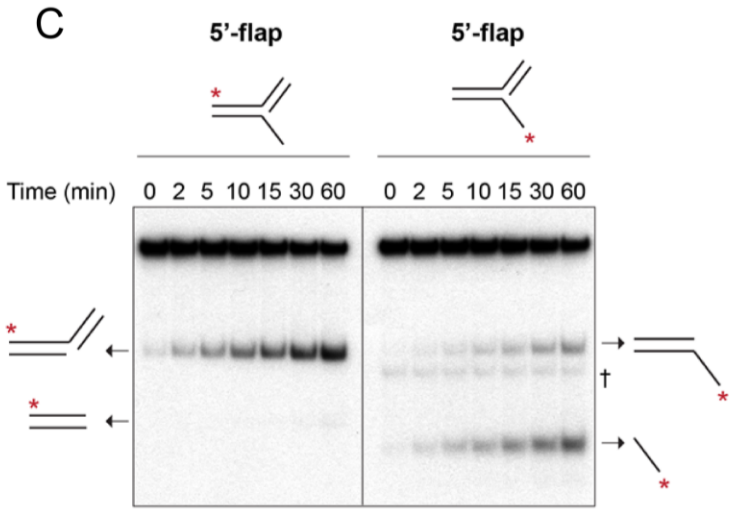
Oligonucleotide	Sequence (5′ – 3′)
X0-1	ACGCTGCCGAATTCTACCAGTGCCTTGCTAGGACATCTTTGCCCACCTGCAGGTTCACCC
X0-2	GGGTGAACCTGCAGGTGGGCAAAGATGTCCATCTGTTGTAATCGTCAAGCTTTATGCCGT
X0-3	ACGGCATAAAGCTTGACGATTACAACAGATCATGGAGCTGTCTAGAGGATCCGACTATCG
X0-4	CGATAGTCGGATCCTCTAGACAGCTCCATGTAGCAAGGCACTGGTAGAATTCGGCAGCGT
X0-1.28	ACGCTGCCGAATTCTACCAGTGCCTTGC
X0-1.30	ACGCTGCCGAATTCTACCAGTGCCTTGCTA
X0-2.30	GGGTGAACCTGCAGGTGGGCAAAGATGTCC
X0-3.30	CATGGAGCTGTCTAGAGGATCCGACTATCG
X0-3.30*	CATGGAGCTGTCTAGAGGATCCGACTATCG
X0-1comp	GGGTGAACCTGCAGGTGGGCAAAGATGTCCTAGCAAGGCACTGGTAGAATTCGGCAGCGT
OligodT_{60}	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT

* This oligonucleotide carried a 5′-phosphate

B

Substrate	Component Oligonucleotides
HJ	X0-1, X0-2, X0-3, X0-4
5′-flap	X0-1, X0-4, X0-2.30
3′-flap	X0-1, X0-4, X0-3.30
Splayed arm	X0-1, X0-4
Gapped duplex	X0-1.28, X0-3.30*, X0-4
Nicked duplex	X0-1.30, X0-3.30*, X0-4
dsDNA	X0-1, X0-1comp
ssDNA	OligodT_{60}

C



D

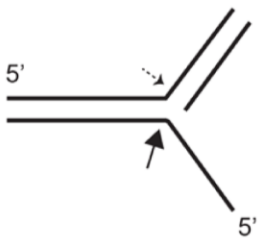


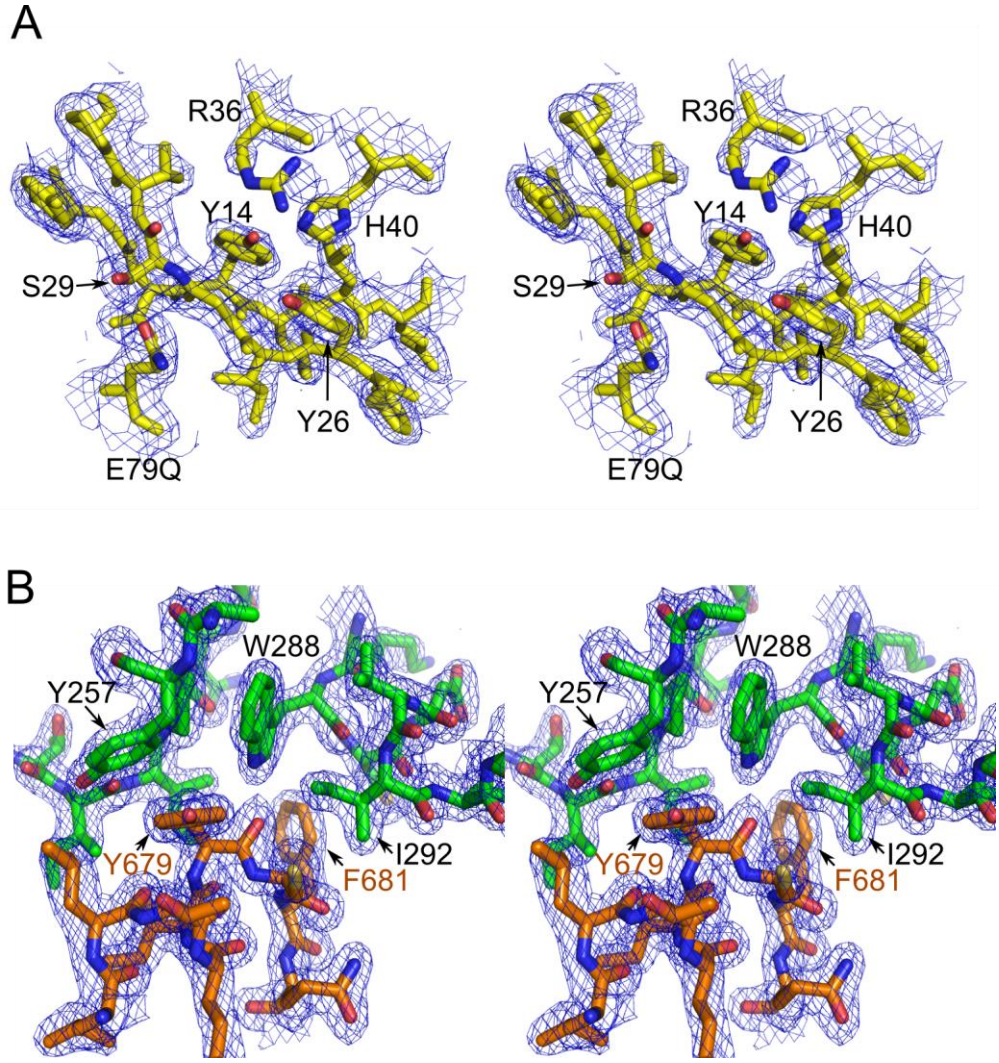
Figure S1. Sequences of DNA oligonucleotides and cleavage of 5'-flap DNA by *Cg*-Slx1-Slx4^{CCD}, related to Figure 1 (previous page).

(A) Sequences of the oligonucleotides used for synthetic substrates.

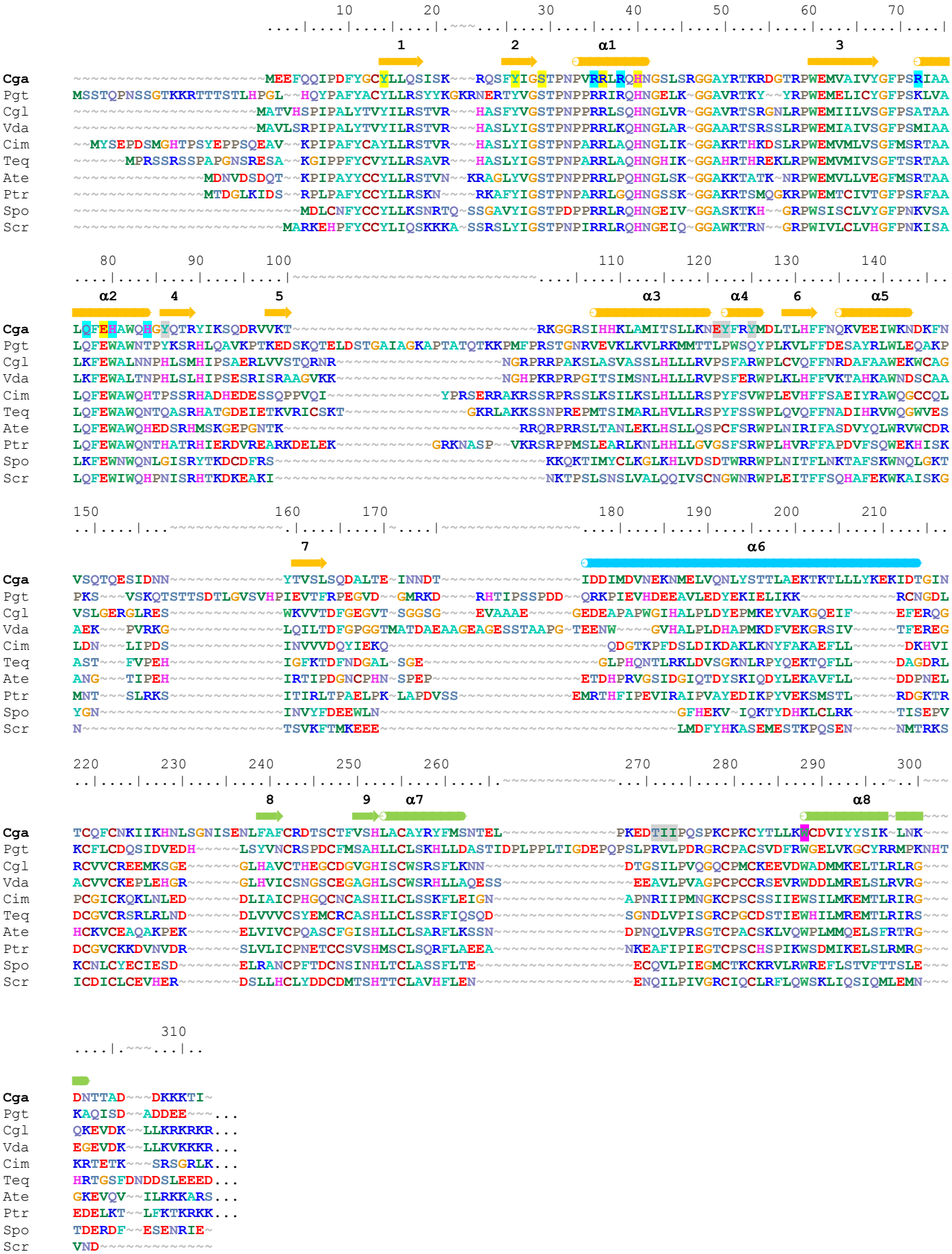
(B) Oligonucleotides annealed to generate synthetic substrates.

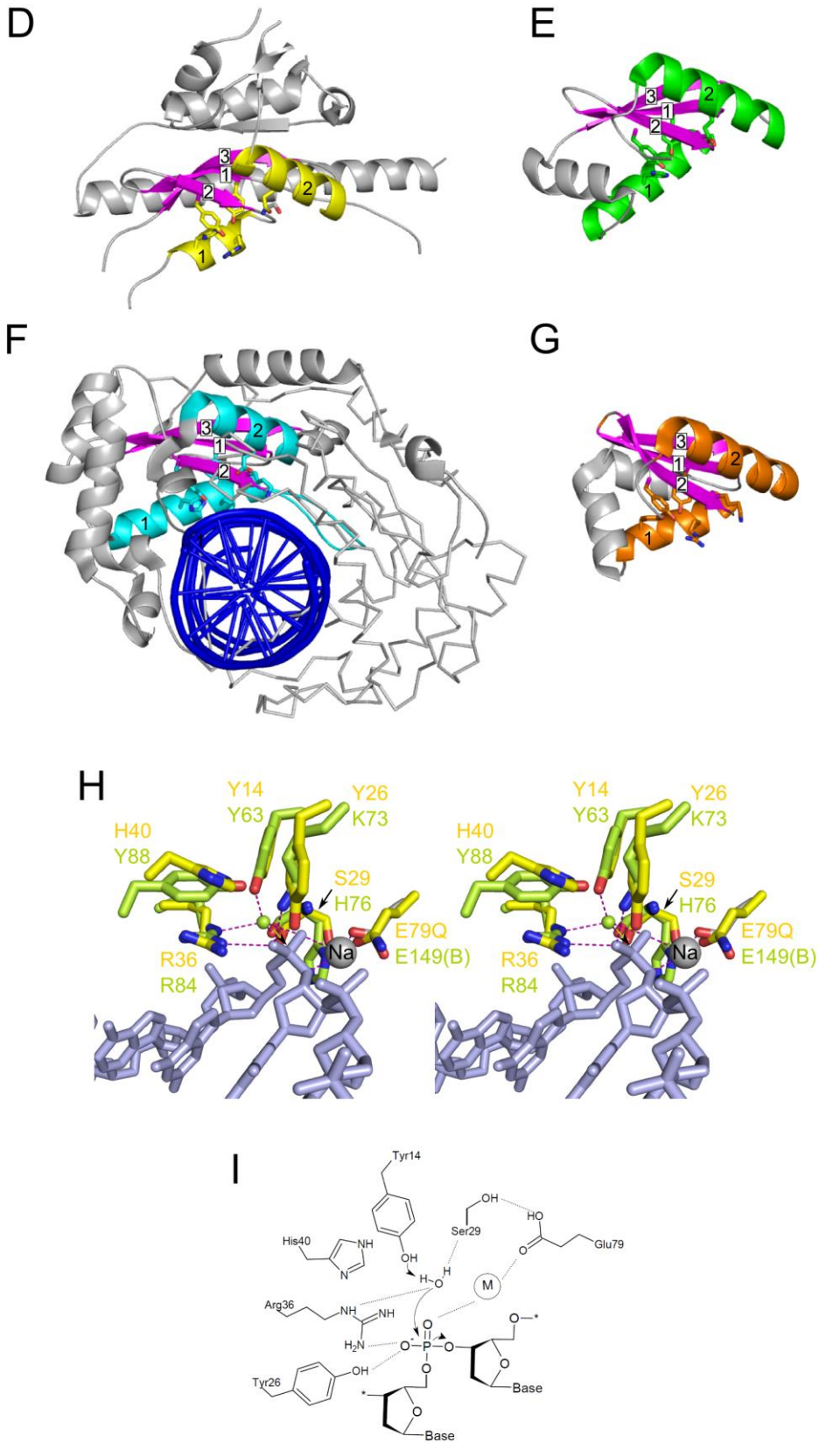
(C) Purified *Cg*-Slx1-Slx4^{CCD} (10 nM) was incubated with 5´-flap DNA (100 nM), spiked with negligible amounts of 5´-^{32}P-labeled substrate, at 37°C for the indicated times. Reaction products were analyzed by native PAGE and autoradiography. The red asterisk denotes the oligonucleotide that was 5'-^{32}P-labeled. The † symbol signifies a DNA product resulting from partial decay of the radiolabelled substrate.

(D) Schematic of the data obtained in (A), depicting the approximate positions of the cleavage sites in the 5'-flap substrates that yield the DNA products shown in (A). The size of each arrow reflects the relative efficiency of cleavage at each site.


A
R36
Y14
H40
S29
Y26
E79Q
R36
Y14
H40
S29
Y26
E79Q
B
W288
Y257
Y679
I292
F681
W288
Y257
Y679
I292
F681

C




D
E
F
G
H
Y14
Y63
Y26
K73
H40
Y88
S29
H76
E79Q
E149(B)
R36
R84
Na
I
Tyr14
Ser29
His40
Glu79
Arg36
Tyr26
Base
M

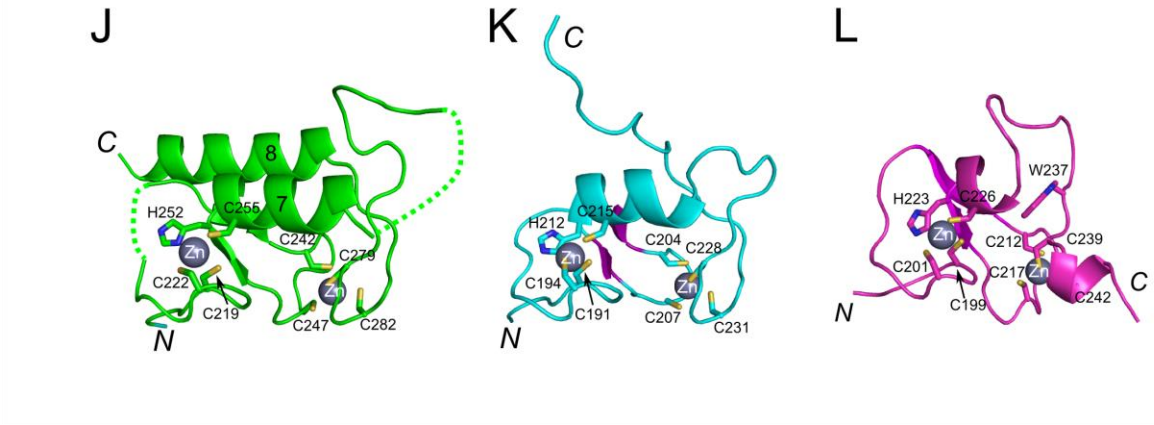


Figure S2. Samples of electron density maps, sequence alignment of fungal Slx1 proteins, comparison of GIY-YIG domains, proposed catalytic mechanism for *Cg*-Slx1-Slx4 and Zinc finger domains (three previous pages), related to Figure 2.

(A) Sample of electron density map of *Cg*-Slx1 structure. A region of the *Cg*-Slx1 active site and fragment of the central β-sheet are shown and a $2F_o$-F_c simulated annealing composite omit map (blue mesh) is overlaid on the structure. The active site residues are labelled.

(B) Sample of electron density map of *Cg*-Slx1-Slx4^{CCD3} structure. A fragment of the interface between Slx1 and Slx4^{CCD3} is shown and a $2F_o$-F_c simulated annealing composite omit map (blue mesh) is overlaid on the structure.

(C) Sequence alignment of fungal Slx1 proteins. Amino acid numbers and secondary structure (tubes for helices and arrows for strands) of *Cg*-Slx1 are given above the sequence alignment. Secondary structure symbols are colored according to the domains shown in Figure 2. Active site residues are highlighted in yellow, predicted DNA-binding amino acids in cyan, residues involved in Slx1 dimerization in gray, and the conserved Slx4-binding tryptophan in magenta. The sequence alignment was performed using Promals3D (Pei et al., 2008). *Candida glabrata* (Cga), *Puccinia graminis tritici* (Pgt), *Chaetomium globosum* (Cgl), *Verticillium dahlia* (Vda), *Coccidioides immitis* (Cim), *Trichophyton equinum* (Teq), *Aspergillus terreus* (Ate), *Pyrenophora tritici-repentis* (Ptr), *Schizosaccharomyces pombe* (Spo), *Schizosaccharomyces cryophilus* (Scr).

(D) Nuclease domain of *Cg*-Slx1.

(E) GIY-YIG domain of homing endonuclease I-TevI (PDB ID: 1LN0 (Van Roey et al., 2002)).

(F) Eco29kI in complex with DNA (PDB ID: 3NIC (Mak et al., 2010)). One protomer of the dimer is shown in cartoon and the other in gray wire representation. The DNA is shown as blue ladder representation.

(G) GIY-YIG domain of UvrC (PDB ID: 1YCZ (Truglio et al., 2005)). In panels (D-G) residues forming the active site are shown as sticks and the conserved secondary structure elements of the GIY-YIG fold (β strands 1-3 and helices α1 and α2) are shown in color and labeled. The non-conserved parts of the structures are shown in gray.

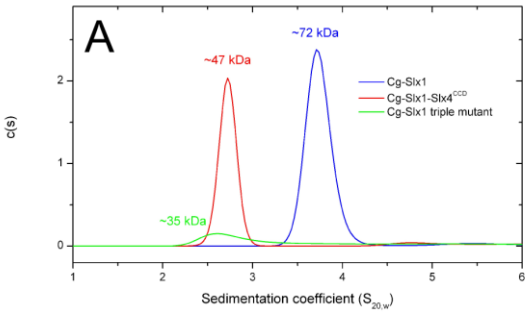
(H) Stereoview of the active site**.** Overlay of the active sites of *Cg*-Slx1 (yellow) and Hpy188I restrictase. The structure of the Hpy188I-DNA complex is shown (PDB ID: 3OQG (Sokolowska et al., 2011)), with the protein colored in green and the DNA substrate colored in blue. The sodium ion is shown as a light gray sphere and the attacking nucleophile as a small green sphere. Interactions of the attacking water and the scissile phosphate, as well as coordination of the metal ion in the Hpy188I structure, are shown as purple dashed lines. The green arrow indicates the direction of the nucleophilic attack. Note that Glu149 in Hpy188I, designated E149(B), comes from the other subunit of the domain-swapped dimer.

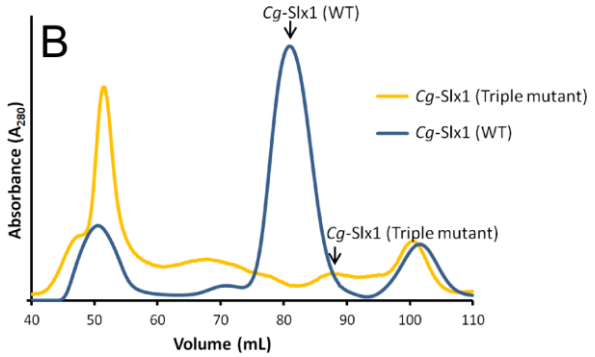
(I) A schematic of the proposed catalytic mechanism.

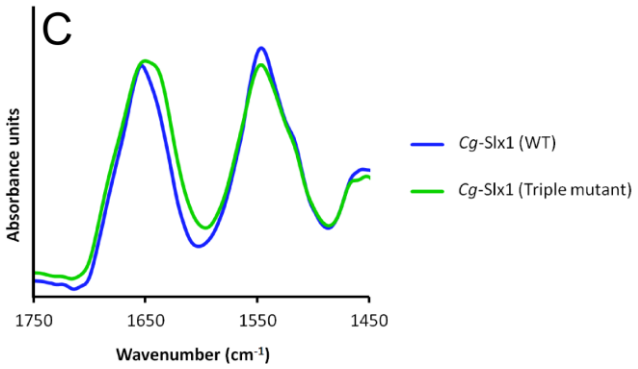
(J) RING domain of *Cg*-Slx1. Zinc ions are shown as gray spheres and coordinating residues as sticks.

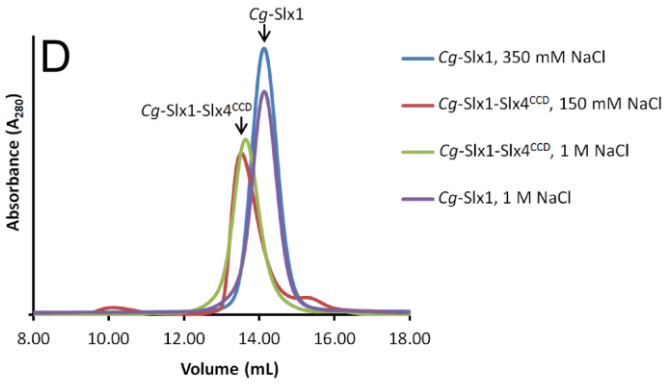
(K) RING domain of NSE1 (PDB ID: 3NW0 (Doyle et al., 2010)).

(L) PHD finger of ING4 (PDB: 2PNX (Hung et al., 2009)). The presence of a tryptophan residue in the PHD finger (shown as sticks) increases the width of the hydrophobic core in comparison to the RING finger domain (Capili et al., 2001).


A
~72 kDa
~47 kDa
~35 kDa
Cg-Slx1
Cg-Slx1-Slx4CCD
Cg-Slx1 triple mutant
c(s)
Sedimentation coefficient (S20,w)


B
Cg-Slx1 (WT)
Cg-Slx1 (Triple mutant)
Cg-Slx1 (WT)
Cg-Slx1 (Triple mutant)
Absorbance (A280)
Volume (mL)


C
Cg-Slx1 (WT)
Cg-Slx1 (Triple mutant)
Absorbance units
Wavenumber (cm-1)


D
Cg-Slx1
Cg-Slx1-Slx4CCD
Cg-Slx1, 350 mM NaCl
Cg-Slx1-Slx4CCD, 150 mM NaCl
Cg-Slx1-Slx4CCD, 1 M NaCl
Cg-Slx1, 1 M NaCl
Absorbance (A280)
Volume (mL)

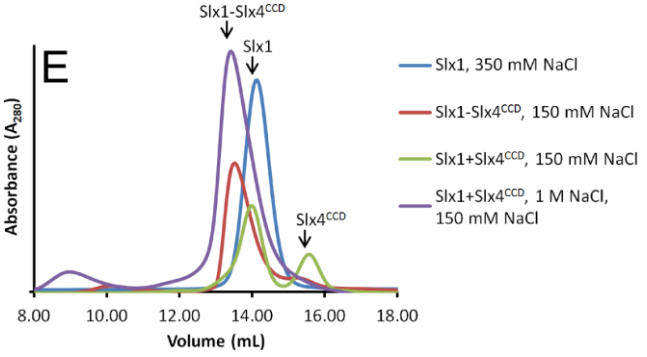

E
Slx1-Slx4CCD
Slx1
Slx4CCD
Slx1, 350 mM NaCl
Slx1-Slx4CCD, 150 mM NaCl
Slx1+Slx4CCD, 150 mM NaCl
Slx1+Slx4CCD, 1 M NaCl,
150 mM NaCl
Absorbance (A280)
Volume (mL)

Figure S3. Oligomeric state of *Cg*-Slx1 and the *Cg*-Slx1-Slx4^{CCD} complex, as determined using analytical ultracentrifugation and gel filtration, and structural stability of dimerization mutants (previous page), related to Figure 3.

(A) Analytical ultracentrifugation, sedimentation velocity analysis. The calculated MW of *Cg*-Slx1-Slx4^{CCD} (red trace) is approximately 47 kDa, consistent with a stable heterodimer containing one molecule each of *Cg*-Slx1 and *Cg*-Slx4^{CCD}. The calculated MW of the *Cg*-Slx1 complex (blue trace) is approximately 72 kDa indicating a stable homodimer. The calculated MW of *Cg*-Slx1$^{Y122A/I272A/I273A}$ (green trace) is approximately 35 kDa, consistent with a monomer.

(B) Purification of *Cg*-Slx1$^{Y122A/I272A/I273A}$ (triple mutant) on HiLoad Superdex 200 PG size exclusion column (yellow trace). *Cg*-Slx1 purification trace is shown for reference (blue). The void volume is at 50 ml. The peaks corresponding to *Cg*-Slx1 (WT) homodimer or monomeric triple mutant protein are shown with arrows.

(C) Fourier-Transform Infrared (FT-IR) spectrum of *Cg*-Slx1 (blue trace) and the monomeric fraction of *Cg*-Slx1$^{Y122A/I272A/I273A}$ (green trace).

(D) Analytical gel filtration analysis. Purified *Cg*-Slx1 and *Cg*-Slx1-Slx4^{CCD} proteins were applied to a Superdex 200 10/300GL size exclusion column in buffer containing 150 mM, 350 mM, or 1 M NaCl.

(E) Reconstitution of the *Cg*-Slx1-Slx4^{CCD} complex by mixing individually purified *Cg*-Slx1 and *Cg*-Slx4^{CCD} proteins in the presence of 1 M NaCl, followed by dialysis against buffer containing 150 mM NaCl and gel filtration.

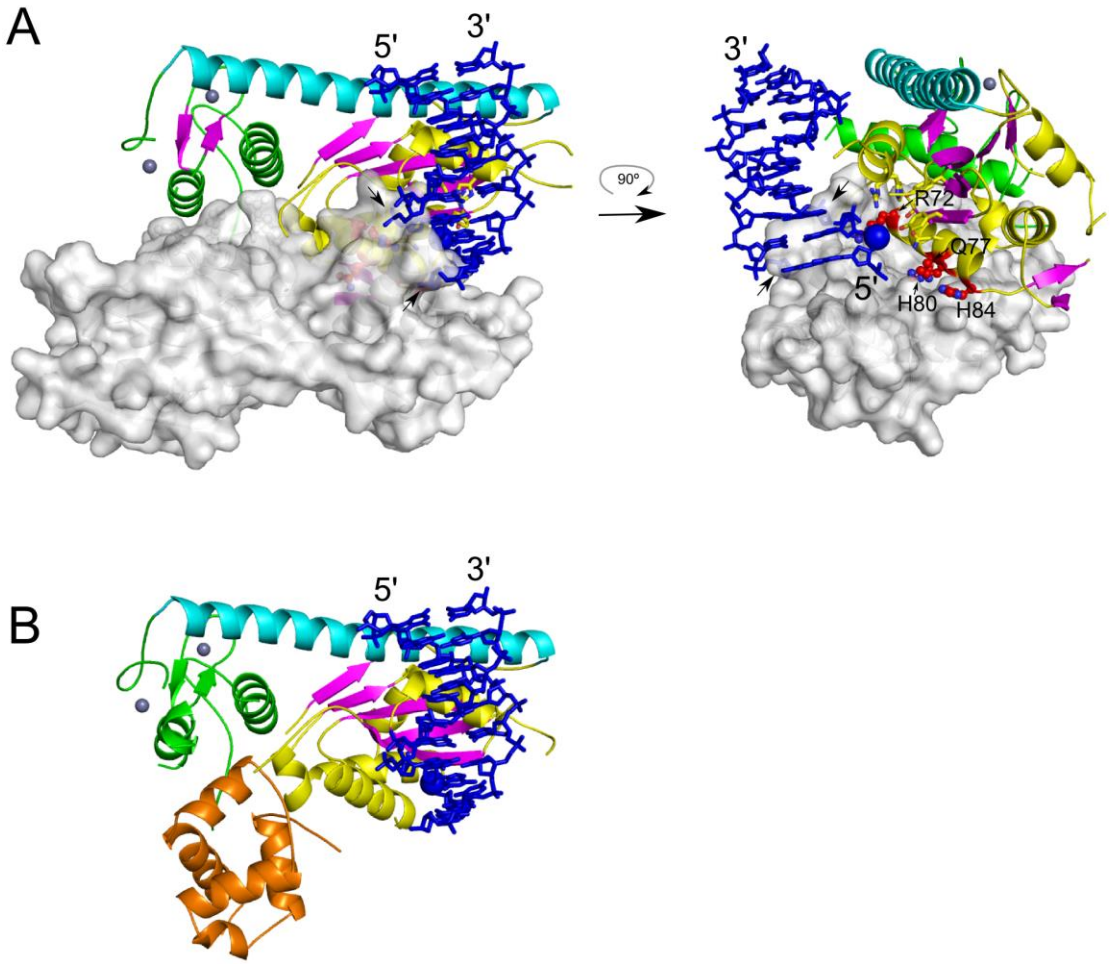


Figure S4. Model of DNA binding by *Cg*-Slx1 homodimer and *Cg*-Slx1-Slx4^{CCD3} complex, related to Figure 4.

(A) Model of *Cg*-Slx1 homodimer interacting with DNA. One subunit of the dimer is shown in cartoon representation, colored as in Figure 2, and the other subunit is shown in white transparent surface. DNA from Eco29kI restrictase-substrate complex (in blue, from PDB: 3NIC (Mak et al., 2010)) is modeled to interact with the active site of the subunit shown in cartoon representation. Only a portion of the dsDNA downstream from the 5'-flap is shown. The blue sphere indicates the scissile phosphate and the active site residues are shown as sticks. Arg72, Asn77, His80 and His84 from helix α2 participate in DNA binding (Figure 4B-D) and are shown in red. Arrows indicate steric clashes between the modeled DNA and the second subunit of the homodimer. Grey spheres represent zinc ions.

(B) Model of the *Cg*-Slx1-Slx4^{CCD3} complex interacting with DNA. The model is shown in the same orientation as in the left panel of (A). *Cg*-Slx1 is colored as in (A) and *Cg*-Slx4^{CCD3} is shown in orange.

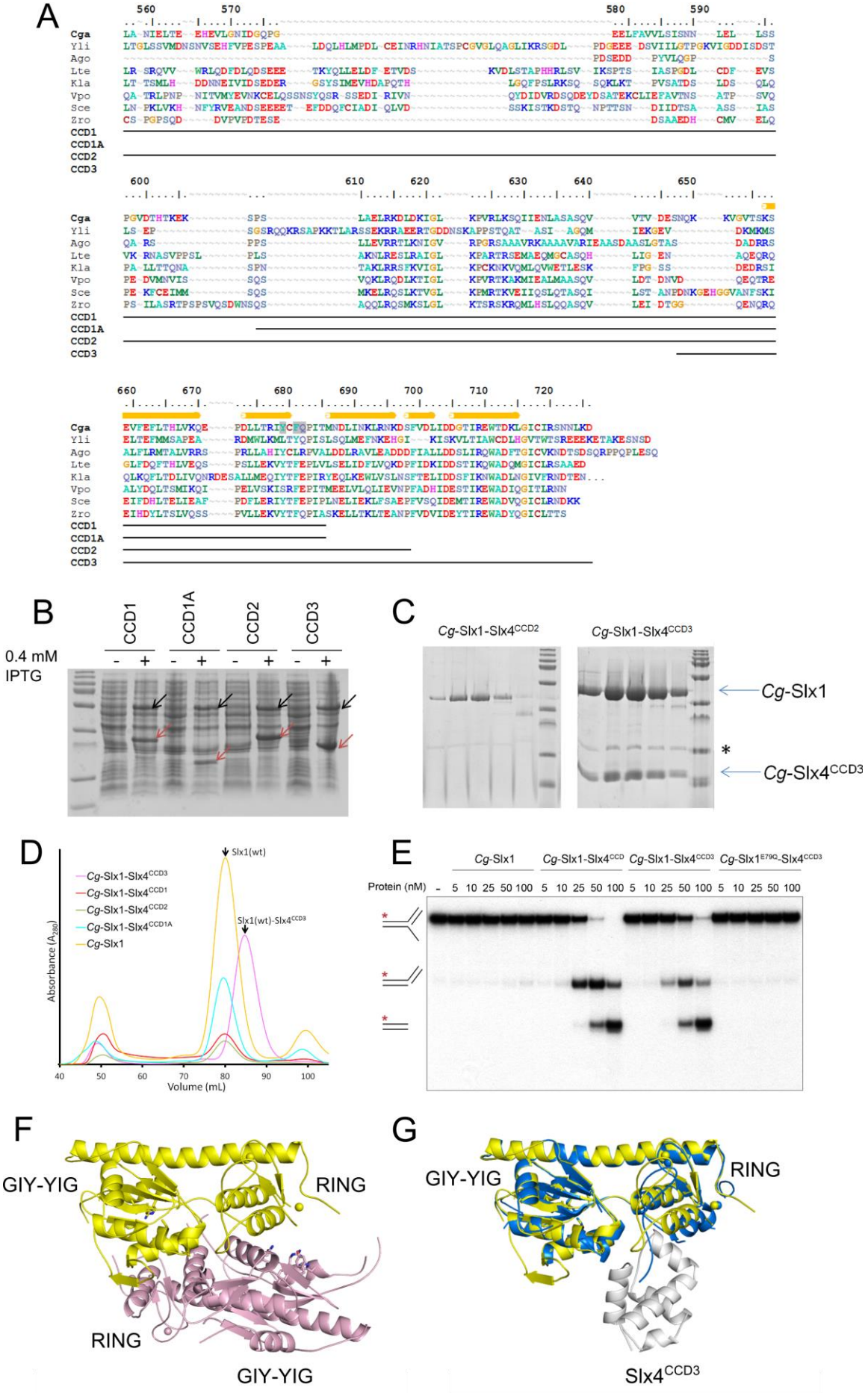

A
560
570
580
590
Cga
Yli
Ago
Lte
Kla
Vpo
Sce
Zro
CCD1
CCD1A
CCD2
CCD3
600
610
620
630
640
650
660
670
680
690
700
710
720
B
CCD1
CCD1A
CCD2
CCD3
0.4 mM IPTG
- + - + - + - +
C
Cg-Slx1-Slx4CCD2
Cg-Slx1-Slx4CCD3
Cg-Slx1
*
Cg-Slx4CCD3
D
Slx1(wt)
Slx1(wt)-Slx4CCD3
Cg-Slx1-Slx4CCD3
Cg-Slx1-Slx4CCD1
Cg-Slx1-Slx4CCD2
Cg-Slx1-Slx4CCD1A
Cg-Slx1
Absorbance (A280)
40 50 60 70 80 90 100
Volume (mL)
E
Cg-Slx1
Cg-Slx1-Slx4CCD
Cg-Slx1-Slx4CCD3
Cg-Slx1E79Q-Slx4CCD3
Protein (nM) - 5 10 25 50 100 5 10 25 50 100 5 10 25 50 100 5 10 25 50 100
F
GIY-YIG
RING
RING
GIY-YIG
G
GIY-YIG
RING
Slx4CCD3

Figure S5. Analysis of *Cg*-Slx1-Slx4 interaction (previous page), related to Figure 5.

(A) Sequence alignment of fungal Slx4 conserved C-terminal domains (CCD). Amino acid numbers of *Cg*-Slx4^{CCD} and secondary structure of *Cg*-Slx4^{CCD3} (tubes denote helices) are given above the sequence alignment. Residues highlighted in gray participate in Slx1-binding. The sequence alignment was performed using Promals3D (Pei et al., 2008). Lines represent the residues present in the four different truncated versions of *Cg*-Slx4^{CCD}. *Candida glabrata* (Cga), *Yarrowia lipolitica* (Yli), *Ashbya gossypii* (Ago), *Lachancea thermotolerans* (Lte), *Kluyveromyces lactis* (Kla), *Vanderwaltozyma polyspora* (Vpo), *Saccharomyces cerevisiae* (Sce), *Zygosaccharomyces rouxii* (Zro).

(B) Over-expression of *Cg*-Slx1 with *Cg*-Slx4^{CCD1-3}. Pre- and post- induction samples (*Cg*-Slx1 and *Cg*-Slx4^{CCD} variants were over-expressed separately and the pellets were mixed prior to lysis), showing the expression of *Cg*-Slx1 and the four *Cg*-Slx4^{CCD} variants. Black arrows indicate *Cg*-Slx1 (fusion with 6xHis-SUMO) and red arrows indicate the *Cg*-Slx4^{CCD} variants (fusion with 6xHis-SUMO).

(C) SDS-PAGE analysis of proteins eluted from a Superdex 200 size exclusion column. The heterodimer was obtained only when *Cg*-Slx1 was co-purified with *Cg*-Slx4^{CCD3}. A contaminating band in the protein purification is marked with an asterisk.

(D) Comparison of elution profiles of *Cg*-Slx1 alone and when co-purified with *Cg*-Slx4^{CCD} variants (HiLoad Superdex 200 PG size exclusion column). The peak at the 50 mL volume corresponds to the void volume of the column.

(E) Nuclease activity assay. The indicated concentrations of purified *Cg*-Slx1-Slx4^{CCD3} were incubated with 5´-flap DNA (100 nM), spiked with negligible amounts of 5´-^{32}P-labeled substrate, at 37°C for 15 min. Reaction products were analyzed by native PAGE and autoradiography. The red asterisk denotes the oligonucleotide that was 5’-^{32}P-labeled.

(F) *Cg*-Slx1 homodimer – one protomer is shown in yellow and the other in pink. Active site residues are shown as sticks. Spheres represent zinc ions.

(G) Superposition of *Cg*-Slx1 and *Cg*-Slx1-Slx4^{CCD3} structures using the GIY-YIG domain to visualize the movement of the RING domain (shown in the same orientation as in (F)). *Cg*-Slx1 (one protomer from the homodimer) is shown in yellow and the *Cg*-Slx1-Slx4^{CCD3} heterodimer in blue (for Slx1) and gray (for Slx4^{CCD3}). Zinc ions are shown as spheres.

SUPPLEMENTAL RESULTS

Comparison of the *Cg*-Slx1 nuclease domain with other GIY-YIG domains

The nuclease domain of *Cg*-Slx1 and other GIY-YIG family members contains a common core fold that is comprised of the central β-sheet strands β1, β2, and β3 and helices α1 and α2 (Figure S2D-G). These enzymes generally contain additional structural features that define their overall architecture and DNA-binding properties. For example, the I-Tev homing endonuclease and UvrC GIY-YIG domain possess one or two additional helices, respectively, inserted in the core of the fold after helix α1 (Figure S2E, G). In Slx1, unique structural elements are located C-terminal to the core of the fold: the β-hairpin comprising strands β4 and β5, the additional strands of the central β-sheet (β6 and β7), and the long helix α6. In the Hpy188I and R.Eco29kI restrictases, multiple additional elements are present but they form very different structures than in Slx1 (Mak et al., 2010; Sokolowska et al., 2011). A distinctive feature of these enzymes is their homodimeric architecture on DNA substrates (Figure S2F). Additionally, domain swapping was observed in the Hpy188I dimer (Sokolowska et al., 2011). Currently, Eco29kI and Hpy188I are the only GIY-YIG nucleases for which crystal structures in complex with DNA substrate are available (Figure S2F).

Active site

Based on the high-resolution structures of substrate and product complexes of the Hpy188I restrictase, a catalytic mechanism for GIY-YIG nucleases has been proposed (Sokolowska et al., 2011). Phosphodiester hydrolysis occurs through an in-line nucleophilic attack of a water molecule or a hydroxide ion on the phosphorus of the scissile phosphate. A single metal ion is bound in the GIY-YIG active site, which is proposed to destabilize the substrate and stabilize the transition state. Relative to the attacking nucleophile, the metal ion is located on the opposite side of the scissile phosphate. The active site of *Cg*-Slx1 is highly conserved with other GIY-YIG nucleases (Figure S2H) and we therefore propose that the catalytic mechanism of Slx1 will be identical to that described for Hpy188I (Sokolowska et al., 2011). Specifically, *Cg*-Slx1 active site residues Tyr14 and Arg36 (Tyr63 and Arg84 in Hpy188I) are predicted to coordinate the attacking water molecule, with Tyr14 performing its deprotonation to form the hydroxide ion (Figure S2H, I). Additionally, the guanidinium group of Arg36 is expected to bind the non-bridging oxygen atom of the scissile phosphate. This oxygen atom would also be stabilized by contacts with the phenolic oxygen of Tyr26 (Lys73 in Hpy188I). *Cg*-Slx1 Glu79 is predicted to coordinate the metal ion, and its replacement with glutamine completely abolished the catalytic activity of *Cg*-Slx1-Slx4^{CCD} (Figure 1B). Two additional, less conserved residues within the Slx1 active site are Ser29 and His40 (Figure S2H, I). Ser29 stabilizes the conformation of Glu79 and its backbone carbonyl likely coordinates the attacking water molecule. In Hpy188I, a tyrosine residue corresponding to *Cg*-Slx1 His40

was postulated to shuttle a hydrogen atom during deprotonation of the attacking nucleophilic water. In summary, the conservation of the active site implies that the catalytic mechanism of Slx1 is the same as that proposed for Hpy188I. Furthermore, the geometry of DNA binding at and around the active site is predicted to be similar between *Cg*-Slx1 and GIY-YIG restrictases, for which DNA-bound structures are known.

RING domain

The C-terminal domain of *Cg*-Slx1 binds two zinc ions, which are coordinated by conserved amino acids. Specifically, His252, Cys219, Cys222, and Cys255 bind one zinc ion and Cys242, Cys247, Cys279, and Cys282 coordinate the second ion (Figure S2J). This domain has been previously classified either as a PHD (Fricke and Brill, 2003) or as a RING finger-like domain (Coulon et al., 2004). In *Cg*-Slx1 the overall structure of this module and the zinc coordination sequence corresponds to a typical C4HC3 type RING finger motif found in E3 ubiquitin ligases (Metzger et al., 2014). When close structural homologs of this *Cg*-Slx1 domain are searched using the DALI server (Holm and Rosenstrom, 2010), the closest hits are those for other RING finger domains of ubiquitin ligases. Visual inspection of those hits identified two very similar RING structures: a human E3 ligase (PDB ID: 3LRQ, Northeast Structural Genomics Consortium) and non-structural maintenance of chromosomes element 1 (NSE1) homolog (PDB ID: 3NW0) (Figure S2K), which associates with melanoma antigen (MAGE), a protein that is highly expressed in tumors (Doyle et al., 2010). The RING domains of *Cg*-Slx1 and NSE1 can be superimposed with a root-mean square deviation (rmsd) of 0.88 Å of the position of 31 pairs of C-α atoms. Moreover, one analysis suggested that a characteristic feature of a PHD finger is a conserved tryptophan residue (Capili et al., 2001), which is not present in the *Cg*-Slx1 zinc finger (Figure S2L). Our structure therefore confirms that the C-terminal part of *Cg*-Slx1 contains a canonical RING domain configuration and not a PHD domain.

Slx1 dimerization

To further study *Cg*-Slx1 homodimerization, we prepared *Cg*-Slx1 variants with substitutions in the dimerization interface: Y122A, Y125A (GIY-YIG domain), T271A, I272A, I273A, and I272A/I273A (RING finger). These proteins were expressed in *E. coli* and purified on a nickel column, but were found to be aggregated when analyzed by gel filtration (not shown). We then prepared *Cg*-Slx1 variants with triple substitutions: Y122A/I272A/I273A, Y125A/I272A/I273A, and T271A/I272A/I273A. Again, these proteins could be expressed in bacteria but were aggregated when analyzed by gel filtration. However, a small peak corresponding to monodisperse protein was observed for the Y122A/I272A/I273A mutant (Figure S3B). The MW of this species, as determined by AUC analysis, corresponded to an Slx1 monomer (Figure S3A), but the protein was catalytically inactive (not shown). Fourier-Transform Infrared (FT-IR) spectra of wild-type *Cg*-Slx1 and monomeric Y122A/I272A/I273A showed clear differences, in particular the shape

of amide I band (~1650 cm^{-1}), which is determined by the secondary structure of the protein (Figure S3C). This data indicates that the structure of the Y122A/I272A/I273A variant is perturbed. Collectively, these results indicate that the monomeric form of *Cg*-Slx1 is intrinsically unstable.

We next wanted to gain further insights into *Cg*-Slx1 self-association and binding to *Cg*-Slx4^{CCD}. To do this, the heterodimeric *Cg*-Slx1-Slx4^{CCD} complex was prepared by co-purification (see Experimental Procedures). We found that both the *Cg*-Slx1-Slx4^{CCD} complex and the *Cg*-Slx1 homodimer were stable, as determined by gel filtration at 1 M NaCl (Figure S3D). When individually purified *Cg*-Slx1 and *Cg*-Slx4^{CCD} proteins were mixed together in the presence of 150 mM NaCl and analyzed by gel filtration, the *Cg*-Slx1 homodimer and *Cg*-Slx4^{CCD} subunit eluted in separate fractions. However, if the proteins were mixed together in 1 M NaCl, followed by a dialysis against 150 mM NaCl, the *Cg*-Slx1-Slx4^{CCD} heterodimer could be reconstituted (Figure S3E). Therefore, the *Cg*-Slx1 homodimer and *Cg*-Slx1-Slx4^{CCD} heterodimer appear to be mutually exclusive and the exchange from an Slx1 homodimer to Slx1-Slx4 heterodimer is promoted by high salt concentrations. The high salt concentration may stabilize the intrinsically unstable monomeric form of *Cg*-Slx1 during the homodimer-heterodimer exchange. Within the cell, chaperone proteins may fulfill this role.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Protein expression and purification

Synthetic genes for *C. glabrata* Slx1 and the conserved C-terminal domain of Slx4 (Slx4^{CCD}) (residues 557 to 726) (Bio Basic Inc., Canada) were codon optimized for expression in *E. coli* and subcloned into a pET28a vector (Novagen) with an N-terminal 6xHis-SUMO tag. All point substitutions and deletions (for generating shorter fragments of *Cg*-Slx4^{CCD}, including *Cg*-Slx4^{CCD1} (residues 557-685), *Cg*-Slx4^{CCD1A} (residues 608-685), *Cg*-Slx4^{CCD2} (residues 557-698) and *Cg*-Slx4^{CCD3} (residues 647-726)) were introduced by QuikChange® Site-Directed Mutagenesis (Agilent Technologies), according to the manufacturer's instructions. Wild type and mutant *Cg*-Slx1 or *Cg*-Slx1-Slx4^{CCD} proteins were expressed in *E. coli* BL21 (DE3) Rosetta™ (Novagen). For protein expression, cells were grown in LB medium at 37°C, induced with 0.4 mM IPTG at OD_{600} = 0.6-0.9, and grown overnight at 12°C. The cells were harvested by centrifugation.

For purification of *Cg*-Slx1 (wild type and substitution mutants), the cell pellet was re-suspended in lysis buffer (20 mM Tris-HCl (pH 8.5), 500 mM NaCl, 5 mM imidazole, 10% (v/v) glycerol, and 5 mM 2-mercaptoethanol) and lysed by sonication. For purification of *Cg*-Slx1-Slx4^{CCD} (and complexes containing the shorter fragments of *Cg*-Slx4^{CCD}), pellets containing over-expressed His-tagged *Cg*-Slx1 and His-

tagged *Cg*-Slx4^{CCD} or its shorter fragments were re-suspended in lysis buffer and mixed together before sonication. The lysate was clarified by centrifugation at 186 000 g (4°C), and the supernatant was loaded onto a HisTrap™ HP column (GE Healthcare) equilibrated in lysis buffer. Proteins were eluted using a linear gradient of imidazole from 5 mM to 500 mM. Fractions containing His-tagged *Cg*-Slx1 and His-tagged *Cg*-Slx4^{CCD} or its shorter fragments were identified by SDS-PAGE, pooled, diluted 5 times with lysis buffer, and incubated with SUMO protease (expressed and purified from a pET28a vector (Novagen) containing the gene for SUMO protease) overnight at 4°C to remove the His tag. The protein was then loaded onto a HisTrap™ column (GE Healthcare) equilibrated in lysis buffer. *Cg*-Slx1 and *Cg*-Slx4^{CCD} were collected in the unbound fraction and subsequently concentrated using a 10 MWCO Amicon® Ultra Centrifugal Filter Device (Millipore). Proteins were further purified on a Superdex 200 size exclusion column (GE Healthcare) in buffer containing 20 mM Tris-HCl (pH 8.5), 500 mM NaCl, 10% (v/v) glycerol, and 5 mM 2-mercaptoethanol.

Oligomeric state of *Cg*-Slx1 and *Cg*-Slx1-Slx4^{CCD}

The oligomeric states of *Cg*-Slx1 and *Cg*-Slx1-Slx4^{CCD} in solution were analyzed using two independent approaches: analytical ultracentrifugation (sedimentation velocity) and multiangle light scattering (MALS). For MALS analysis, protein samples (500 µL) were applied on a silica size exclusion column (Waters BioSuite 250HR, 10-500kDa) equilibrated with 20 mM HEPES-NaOH pH 7.5, 350 mM NaCl, 5% glycerol and 4 mM 2-mercaptoethanol) with three in-line detectors: UV absorbance, MALS (DAWN HELEOS-II, Wyatt Technology), and differential refractometer (Optilab T-rEX, Wyatt Technology). Data processing and molecular weight calculations were performed using ASTRA software (Wyatt Technology). Sedimentation velocity experiments were performed in a Beckman-Coulter ProteomeLab XL-I analytical ultracentrifuge, equipped with AN-50Ti rotor (8-holes) and 12 mm path length, double-sector charcoal-Epon cells, loaded with 400 µL of samples and 410 µL of buffer (50 mM HEPES-NaOH, pH 7.5, 350 mM NaCl). The experiments were carried out at 4°C and 50,000 rpm, using continuous scan mode and radial spacing of 0.003 cm. Scans were collected in 8 min intervals at 260 nm and 280 nm. The fitting of absorbance versus cell radius data was performed using SEDFIT software, version 14.3e (Schuck, 2000) and continuous sedimentation coefficient distribution c(s) model, covering the range of 0.1–10 S. Biophysical parameters of the buffer (density ρ = 1.01594 g/cm^3 (4°C) and viscosity η = 0.01641 poise (4°C)) and proteins (partial specific volume (V-bar)) were calculated using SEDNTERP software (version 1.09, http://www.jphilo.mailway.com/download.htm). V-bar for *Cg*-Slx1-Slx4^{CCD} was calculated to be 0.7282 cm^3/g (4°C).

DNA substrates for nuclease assays

Synthetic DNA substrates were prepared by annealing partially complementary oligonucleotides (Sigma-Aldrich), as described previously (Wyatt et al., 2013). The oligonucleotide sequences are provided in Figure S1A and those used to construct each DNA substrate are listed in Figure S1B. For non-radiolabeled substrates, 600 pmol 60-mer oligonucleotide(s) and 1200 pmol 30-mer oligonucleotide(s) were mixed in annealing buffer (150 mM NaCl, 15 mM $Na_3C_6H_5O_7$) and incubated in a 95°C water bath for 2 min, after which time the heat was turned off to allow slow cooling to room temperature overnight. On the next day, the annealing reactions were mixed with native DNA loading dye (6X = 30% glycerol, 0.25% w/v bromophenol blue, 0.25% w/v xylene cyanol) and electrophoresed through 12% native polyacrylamide gels for 4-5 hr at 200 V in TBE running buffer (4°C). Following electrophoresis, DNA was visualized using UV shadowing on POLYGRAM® CEL 300 PEI/UV_{254} (Macherey-Nagel) thin layer chromatography paper. Bands representing the fully annealed substrates were excised from the gel and eluted in 500 µL TMgN buffer (10 mM Tris-HCl pH 8.0, 1 mM $MgCl_2$, 50 mM NaCl) overnight at 4°C. Substrate concentration was determined by spectrophotometry at λ = 260 nm. ^{32}P-labeled DNA substrates were prepared as described previously (Rass and West, 2006), with the exception that 10 pmol of oligonucleotide was radiolabelled using 3 µL of [γ-^{32}P]ATP (3000 Ci/mmol, 10 mCi/mL) (GE Healthcare).

Nuclease assays

Unless indicated otherwise, nuclease assays (10 µL) were performed with 10 nM enzyme (or 5, 10, 25, 50, 100 nM for enzyme titrations) and 100 nM non-radiolabelled DNA substrate (spiked with 0.1 µL of ^{32}P-labeled substrate) in cleavage buffer optimized for *Cg*-Slx1-Slx4^{CCD} activity (50 mM Tris-HCl pH 8.5, 2.0 mM $MgCl_2$, 0.1 mg/mL BSA, 1 mM DTT; data not shown). Reactions were assembled without protein and equilibrated at 37°C for 10 min. Cleavage reactions were initiated by the addition of purified protein, incubated at 37°C for the indicated times, and quenched in stop buffer (1 mM Tris-HCl pH 8.0, 2 mg/mL proteinase K (Promega), 0.1% SDS, 2 mM $CaCl_2$) for 30-60 min at 37°C. Terminated reactions were supplemented with native DNA loading dye (6X = 30% glycerol, 0.25% w/v bromophenol blue, 0.25% w/v xylene cyanol) and electrophoresed through 10% native polyacrylamide gels for 75 min at 150 V in TBE running buffer. Gels were dried on Whatman DE81 chromatography paper and analyzed by autoradiography. The results were quantified by phosphoimaging using a Typhoon scanner and ImageQuant® software (GE Healthcare). The cleavage products are expressed as the mean percentage of total radiolabeled DNA ± SEM, calculated from at least three independent experiments.

Electrophoretic mobility shift assays

Synthetic 5'-flap DNA substrates were prepared as described above, with the exception that 50 pmol of oligonucleotide X0-1 was labeled using 1.0 µl of [γ-^{33}P]ATP (3000 Ci/mmol, 10 mCi/mL) (Haartmann Analytic) in a reaction catalyzed by T4 polynucleotide kinase (Thermo Scientific).

Electrophoretic mobility shift assays were performed using the indicated enzyme concentration, 105 nM 5´-flap DNA and 20 nM ^{33}P-labeled substrate. Reaction mixtures were incubated at room temperature for 20 min to allow equilibration before subjecting them to electrophoresis using a 6% TBE polyacrylamide gel. The gel was run at 100 V in TBE buffer. Gels were dried on Whatman DE81 chromatography paper and analyzed by autoradiography. The results were quantified by phosphoimaging using a Typhoon scanner and ImageQuant® software (GE Healthcare). The bound substrate is expressed as the mean percentage of total bound radiolabeled DNA ± SEM, calculated from at least three independent experiments.

Fourier Transform – Infrared Spectroscopy

Secondary structure analysis was carried out using Fourier Transform – Infrared (FT-IR) spectroscopy using a Bruker Tensor 27 FT-IR spectrometer. Protein samples were used at a concentration of 1 mg ml^{-1} in 20 mM HEPES-NaOH (pH 7.5) and 350 mM NaCl. Results were analyzed using OPUS-PRO software.

SUPPLEMENTAL REFERENCES

Capili, A.D., Schultz, D.C., Rauscher, I.F., and Borden, K.L. (2001). Solution structure of the PHD domain from the KAP-1 corepressor: structural determinants for PHD, RING and LIM zinc-binding domains. The EMBO Journal *20*, 165-177.

Coulon, S., Gaillard, P.H., Chahwan, C., McDonald, W.H., Yates, J.R., 3rd, and Russell, P. (2004). Slx1-Slx4 are subunits of a structure-specific endonuclease that maintains ribosomal DNA in fission yeast. Molecular Biology of the Cell *15*, 71-80.

Doyle, J.M., Gao, J., Wang, J., Yang, M., and Potts, P.R. (2010). MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases. Molecular Cell *39*, 963-974.

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

Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Research *38*, W545-549.

Hung, T., Binda, O., Champagne, K.S., Kuo, A.J., Johnson, K., Chang, H.Y., Simon, M.D., Kutateladze, T.G., and Gozani, O. (2009). ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. Molecular Cell *33*, 248-256.

Mak, A.N., Lambert, A.R., and Stoddard, B.L. (2010). Folding, DNA recognition, and function of GIY-YIG endonucleases: crystal structures of R.Eco29kI. Structure *18*, 1321-1331.

Metzger, M.B., Pruneda, J.N., Klevit, R.E., and Weissman, A.M. (2014). RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. Biochimica et Biophysica Acta *1843*, 47-60.

Pei, J., Kim, B.H., and Grishin, N.V. (2008). PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Research *36*, 2295-2300.

Rass, U., and West, S.C. (2006). Synthetic junctions as tools to identify and characterize Holliday junction resolvases. Methods in Enzymology *408*, 485-501.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophysical Journal *78*, 1606-1619.

Sokolowska, M., Czapinska, H., and Bochtler, M. (2011). Hpy188I-DNA pre- and post-cleavage complexes--snapshots of the GIY-YIG nuclease mediated catalysis. Nucleic Acids Research *39*, 1554-1564.

Truglio, J.J., Rhau, B., Croteau, D.L., Wang, L., Skorvaga, M., Karakas, E., DellaVecchia, M.J., Wang, H., Van Houten, B., and Kisker, C. (2005). Structural insights into the first incision reaction during nucleotide excision repair. The EMBO Journal *24*, 885-894.

Van Roey, P., Meehan, L., Kowalski, J.C., Belfort, M., and Derbyshire, V. (2002). Catalytic domain structure and hypothesis for function of GIY-YIG intron endonuclease I-TevI. Nature Structural Biology *9*, 806-811.

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