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

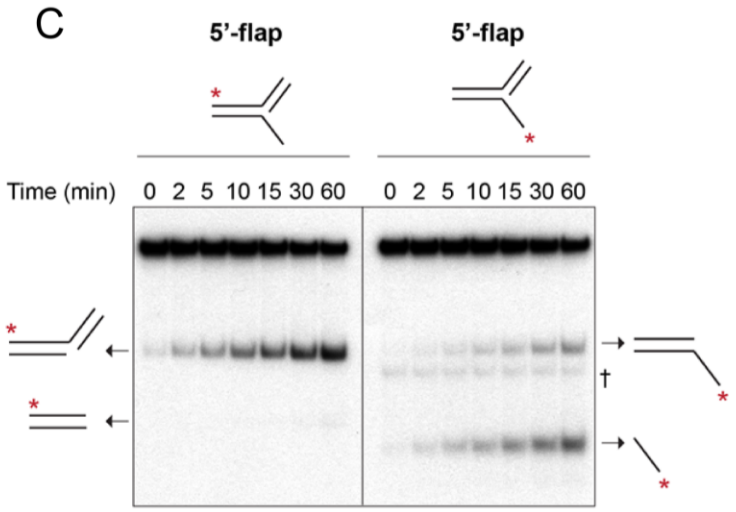
Oligonucleotide	Sequence (5′ – 3′)
X0-1	ACGCTGCCGAATTCTACCAGTGCCTTGCTAGGACATCTTTGCCCACCTGCAGGTTCACCC
X0-2	GGGTGAACCTGCAGGTGGGCAAAGATGTCCATCTGTTGTAATCGTCAAGCTTTATGCCGT
X0-3	ACGGCATAAAGCTTGACGATTACAACAGATCATGGAGCTGTCTAGAGGATCCGACTATCG
X0-4	CGATAGTCGGATCCTCTAGACAGCTCCATGTAGCAAGGCACTGGTAGAATTCGGCAGCGT
X0-1.28	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

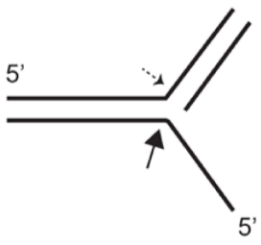


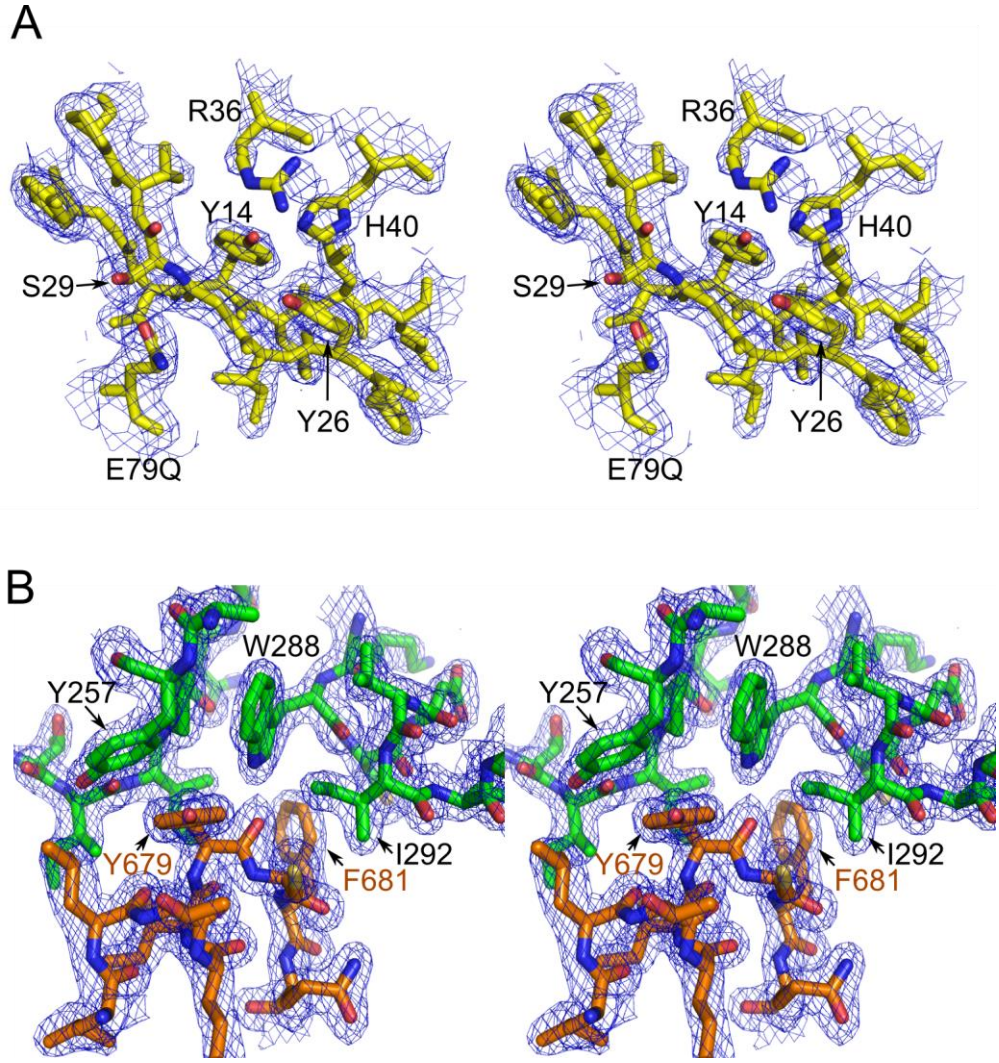
Figure S1. Sequences of DNA oligonucleotides and cleavage of 5'-flap DNA by *Cg*-Slx1-Slx4^{CCD}, related to Figure 1 (previous page).

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

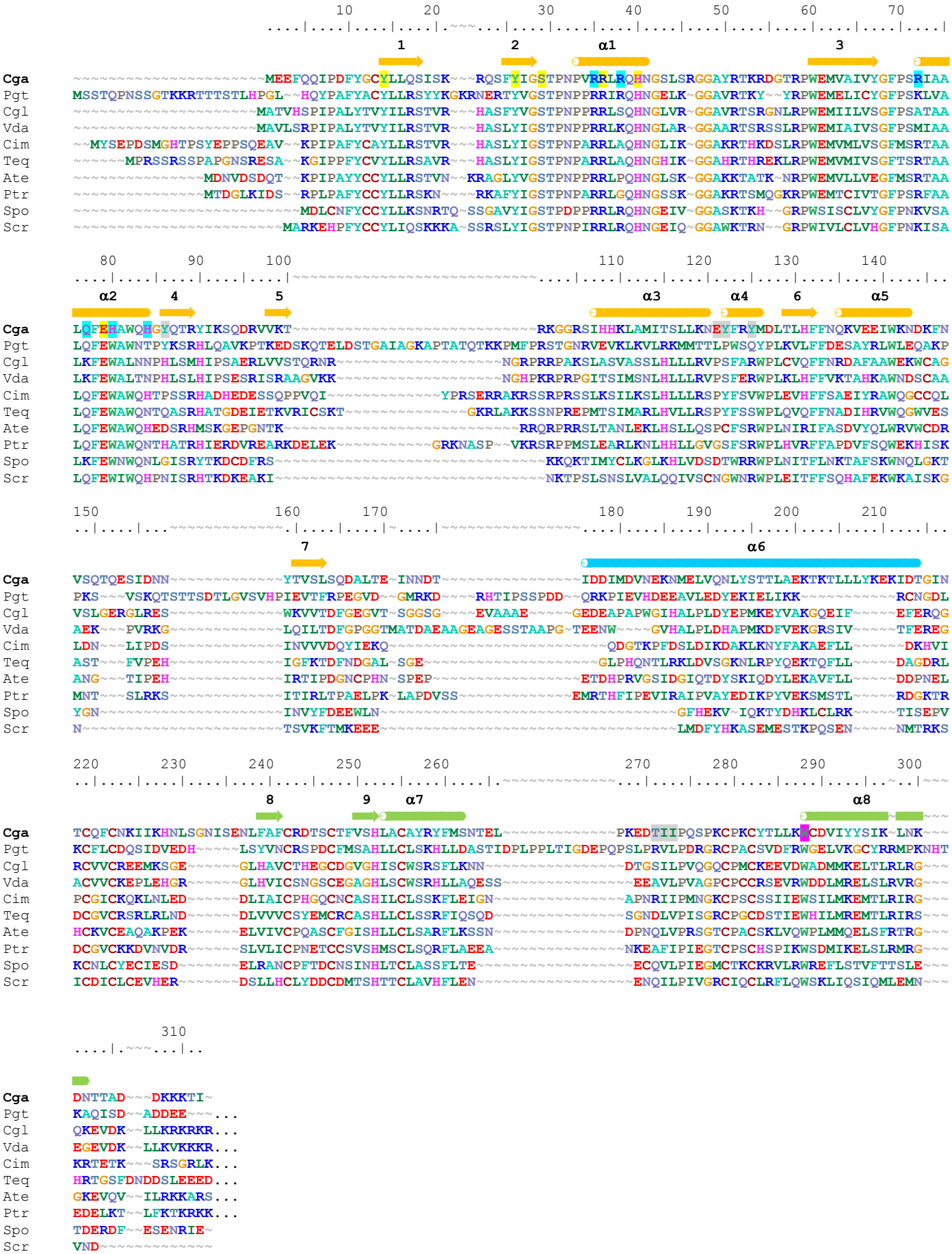
(B) Oligonucleotides annealed to generate synthetic substrates.

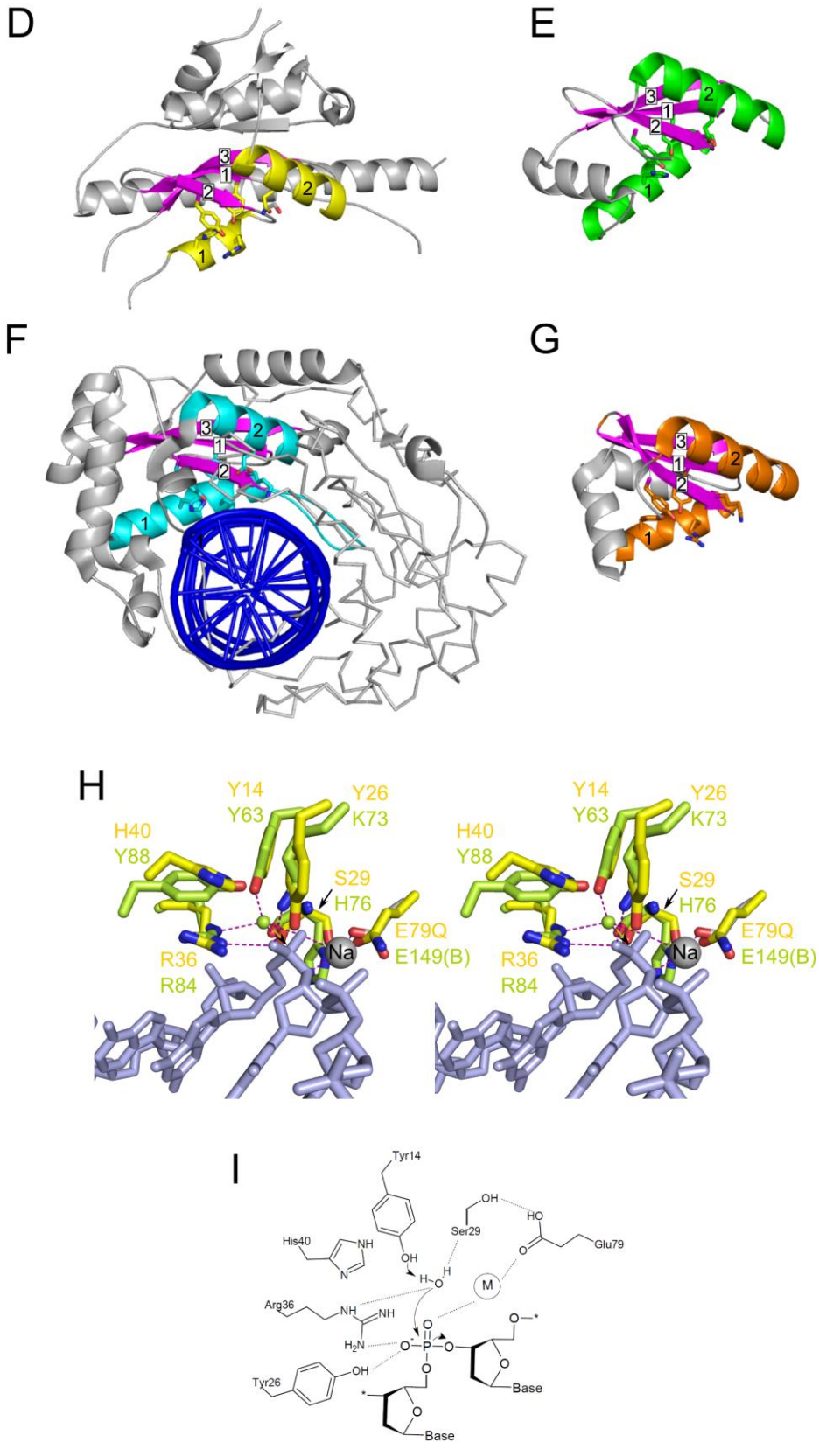
(C) Purified *Cg*-Slx1-Slx4^{CCD} (10 nM) was incubated with 5´-flap DNA (100 nM), spiked with negligible amounts of 5´-^{32}P-labeled substrate, at 37°C for the indicated times. Reaction products were analyzed by native PAGE and autoradiography. The red asterisk denotes the oligonucleotide that was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

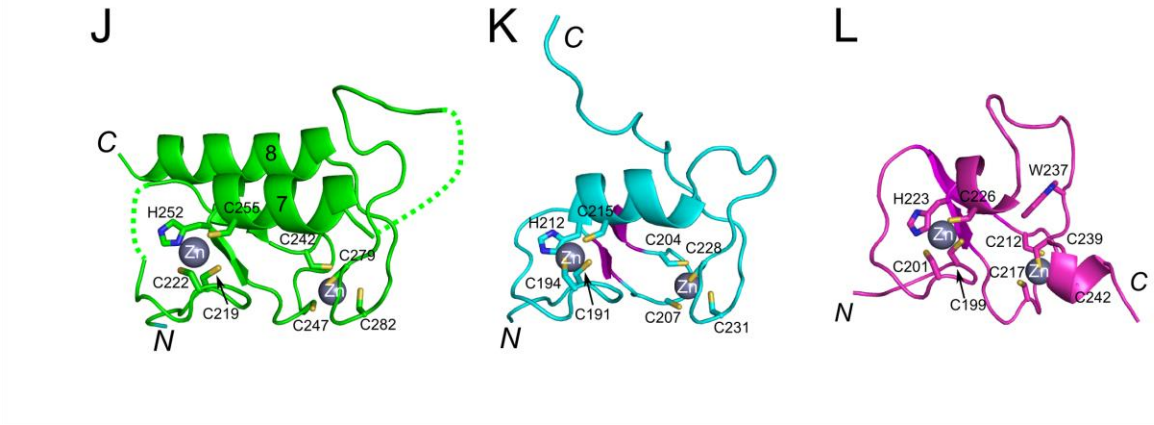


Figure S2. Samples of electron density maps, sequence alignment of fungal Slx1 proteins, comparison of GIY-YIG domains, proposed catalytic mechanism for *Cg*-Slx1-Slx4 and Zinc finger domains (three previous pages), related to Figure 2.

(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.

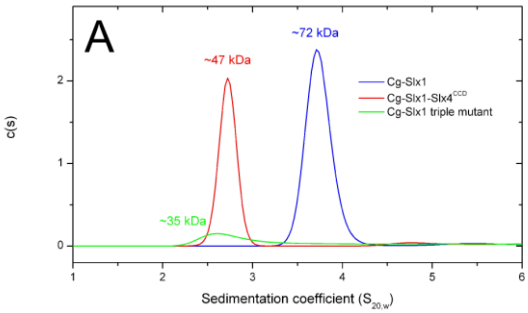
(H) Stereoview of the active site**.** Overlay of the active sites of *Cg*-Slx1 (yellow) and Hpy188I restrictase. The structure of the Hpy188I-DNA complex is shown (PDB ID: 3OQG (Sokolowska et al., 2011)), with the protein colored in green and the DNA substrate colored in blue. The sodium ion is shown 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.

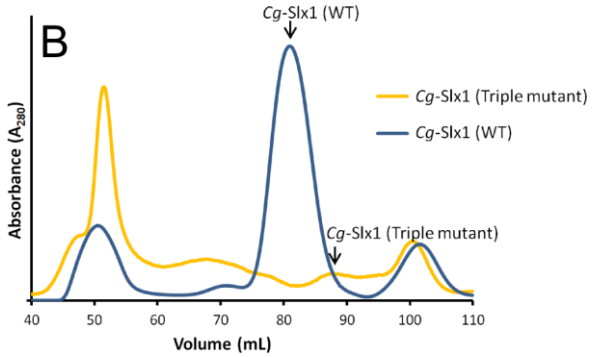
(I) A schematic of the proposed catalytic mechanism.

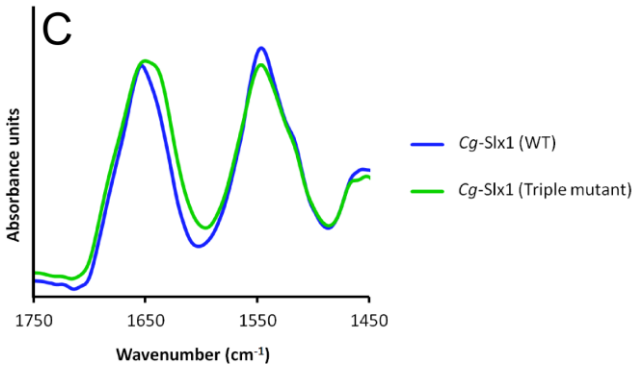
(J) RING domain of *Cg*-Slx1. Zinc ions are shown as gray spheres and coordinating residues as sticks.

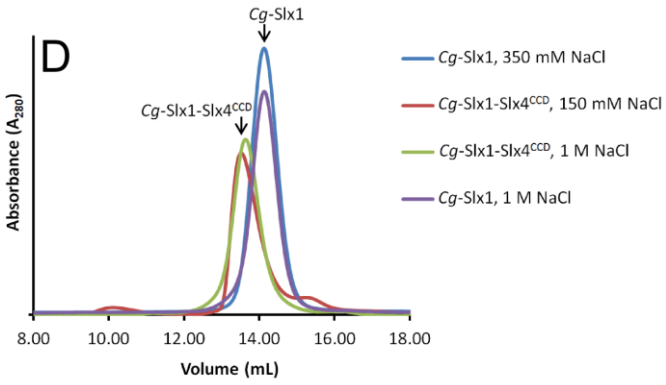
(K) RING domain of NSE1 (PDB ID: 3NW0 (Doyle et al., 2010)).

(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)

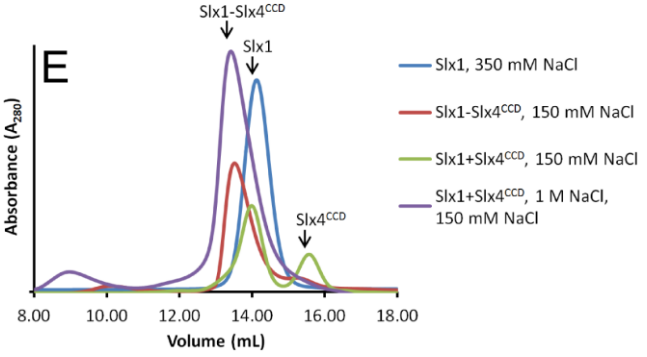

E
Slx1-Slx4CCD
Slx1
Slx4CCD
Slx1, 350 mM NaCl
Slx1-Slx4CCD, 150 mM NaCl
Slx1+Slx4CCD, 150 mM NaCl
Slx1+Slx4CCD, 1 M NaCl,
150 mM NaCl
Absorbance (A280)
Volume (mL)

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.

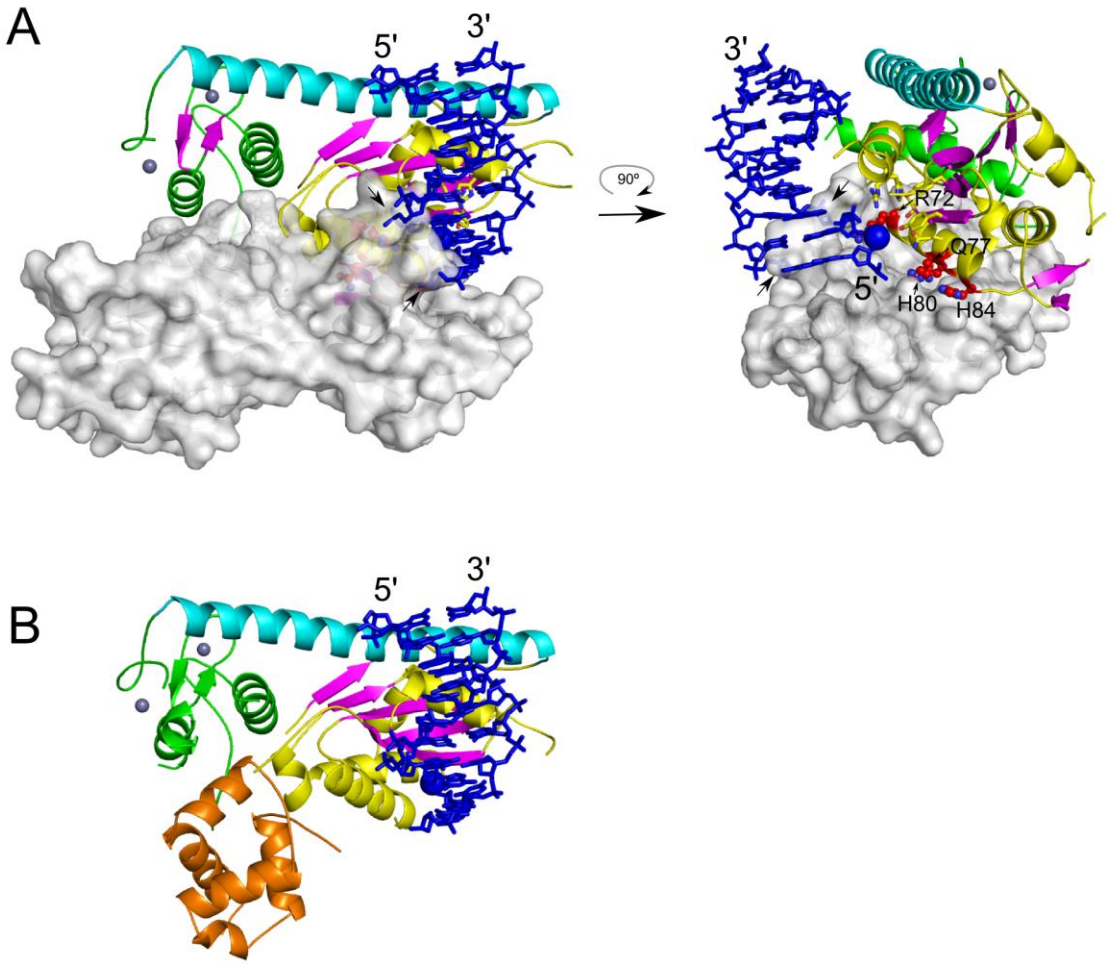


Figure S4. Model of DNA binding by *Cg*-Slx1 homodimer and *Cg*-Slx1-Slx4^{CCD3} complex, related to Figure 4.

(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.

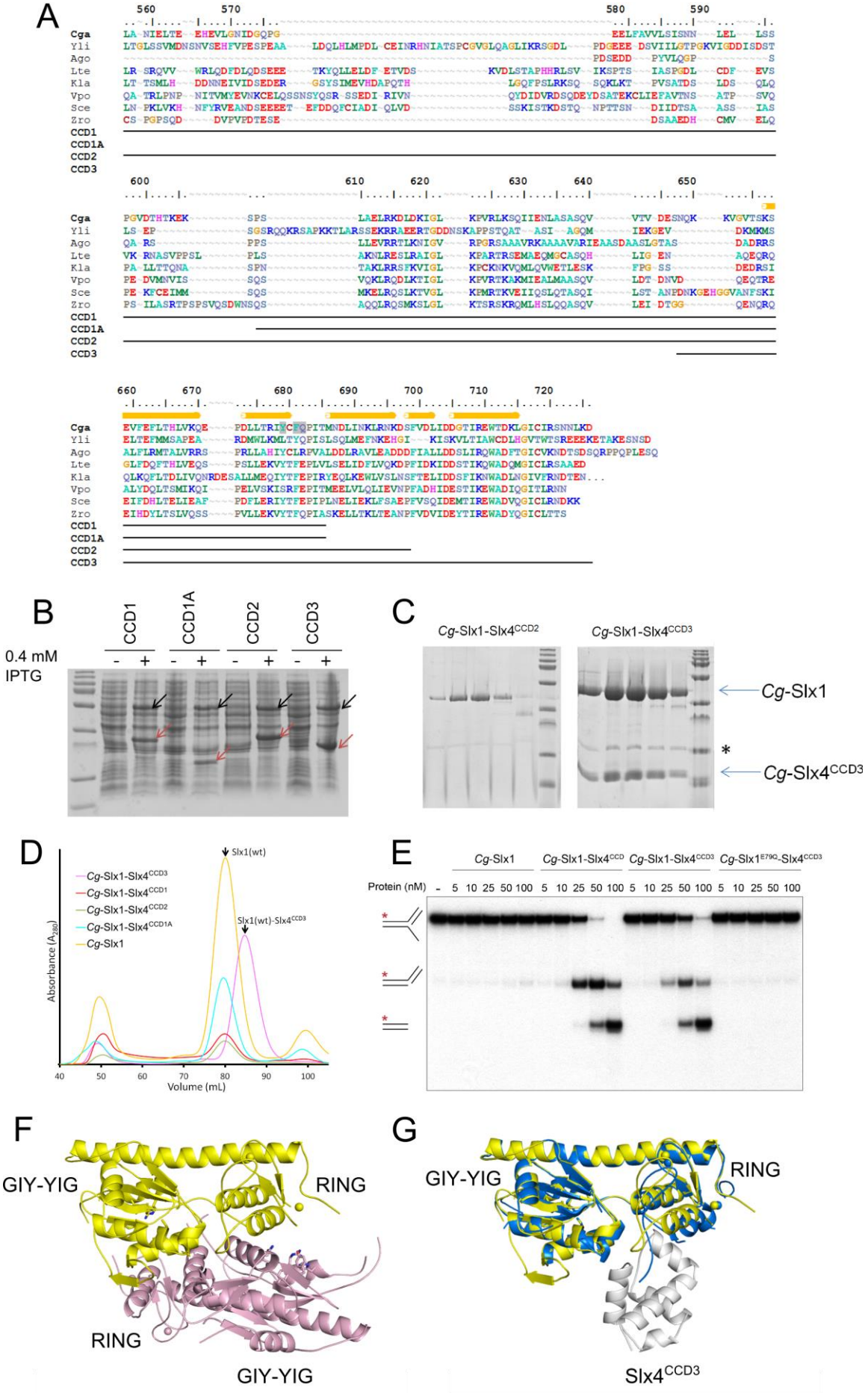

A
560
570
580
590
Cga
Yli
Ago
Lte
Kla
Vpo
Sce
Zro
CCD1
CCD1A
CCD2
CCD3
600
610
620
630
640
650
660
670
680
690
700
710
720
B
CCD1
CCD1A
CCD2
CCD3
0.4 mM IPTG
- + - + - + - +
C
Cg-Slx1-Slx4CCD2
Cg-Slx1-Slx4CCD3
Cg-Slx1
*
Cg-Slx4CCD3
D
Slx1(wt)
Slx1(wt)-Slx4CCD3
Cg-Slx1-Slx4CCD3
Cg-Slx1-Slx4CCD1
Cg-Slx1-Slx4CCD2
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.