Supplementary information for

SLX4 dampens MutSα-dependent mismatch repair

Jean-Hugues Guervilly1*, Marion Blin1†, Luisa Laureti1†, Emilie Baudelet1, Stéphane Audebert1, Pierre-Henri Gaillard1*

1Centre de Recherche en Cancérologie de Marseille, CRCM, Inserm, CNRS, Aix-Marseille Université, Institut Paoli-Calmettes, Marseille, France

* To whom correspondence should be addressed. Tel: +33 486 97 73 71; Fax: +33 486 97 74 99; Email: pierre-henri.gaillard@inserm.fr

Correspondence may also be addressed to: jean-hugues.guervilly@inserm.fr

† These authors contributed equally to this work

This pdf includes:

- Supplementary results (related to Figure 2 and Figure S2 and S3)
- Supplementary methods: mass spectrometry analysis
- Supplementary figure legends
- Supplementary figures:
  - Figure S1 (related to Figure 1)
  - Figure S2 (related to Figure 2)
  - Figure S3 (related to Figure 2)
  - Figure S4 (related to Figure 3 and Figure 4)
  - Figure S5 (related to Figure 5)

- Supplementary references
SUPPLEMENTARY RESULTS

Molecular characterization of SLX4ΔNter in KO30 cells

We selected the HeLa KO clone number 30 (KO30) for further studies based on the initial apparent absence of SLX4 expression as judged by WB (Fig S2D). However, a band of lower molecular weight (MW) became detectable by a C-terminal specific SLX4 antibody at later passages of HeLa KO30, as in other clones (Fig 2A, Fig S2C and S2D). This suggested that cells producing an N-terminally truncated SLX4 protein (SLX4ΔNter), which might confer a growth advantage, are selected through successive passages (Fig 2A). Importantly, the SLX4ΔNter band was lost following siRNA-mediated depletion of SLX4 (Fig 2A) while it was specifically detected in SLX4 immunoprecipitates (IPs) from HeLa KO30 cell extracts (Fig S3A). Of note, an aspecific band that runs at the same position than SLX4ΔNter was sometimes detected in whole cell extracts of parental HeLa FITo cells but not in SLX4 IPs (Fig S3A). The N-terminally truncated nature of the SLX4 specie detected in HeLa KO30 cells was further confirmed by IP/MS analyses of SLX4 IPs. In contrast to SLX4 IPs from the parental FITo, SLX4 IPs from HeLa KO30 cells were systematically devoid of peptides spanning the N-terminus of SLX4 (Fig S3B).

PCR analysis of genomic DNA from HeLa KO30 cells revealed the disruption of SLX4 exon 3 by the insertion of the HR-plasmid (Fig S3C and S3D). The production of SLX4ΔNter in these cells suggested that an alternative SLX4 mRNA that contains an alternative Translation Initiation Site (TIS) is transcribed from a start site located downstream of the plasmid insertion site. In line with this, we noticed that several alternative Transcription Start Sites (TSS) have been identified within intron 3 and at the beginning of exon 4 using TSS-seq (DataBase of Transcriptional Start Sites https://dbtss.hgc.jp/) in various cell lines (Fig S3E). Furthermore, our in silico analyses of sequences located downstream of these alternative TSS, identified Methionine 360 (Met360) as a good candidate for an alternative TIS (altTIS) that matches the Kozak sequence consensus (Fig S3F) (1). However, Met360 is the first aa of a tryptic peptide (MEVGQPQLLQAVR) that we reproducibly detected (n=14) in SLX4 IPs from HeLa FITo cells but never in SLX4 IPs from HeLa KO30 cells (Fig S3G). We surmised that this might be due to a post-translational modification (PTM) of Met360 in HeLa KO30 cells that blurs MS/MS analysis and peptide identification. In line with this, an estimated 97,5% of initiator Methionines (iMet) that are immediately followed by a Glutamate residue, as is Met360, are N-terminally acetylated (2). To establish whether Met360 might be acetylated in HeLa KO30 cells but not in HeLa FITo parental cells, we took into account for peptide mass calculation the possible acetylation of Methionine residues in IP/MS analyses of SLX4 IPs from nuclear extracts derived from both cell lines. As shown in Figure S3H, this enabled us to detect a Met-acetylated peptide (Ac-MEVGQPQLLQAVR) specifically in SLX4 IPs from HeLa KO30 cells, which demonstrates that Met360 serves as the iMet of SLX4ΔNter. Our data therefore unambiguously show that SLX4ΔNter is a shorter version of SLX4 (aa360-1834) produced in HeLa KO30 cells that lacks the MSH2-binding SHIP box, identified in this study, and the well characterized tandem UBZ4 ubiquitin binding zinc fingers (Fig 2B).
**SUPPLEMENTARY METHODS**

**Mass spectrometry analysis**

*Interactome analysis*

Immunoprecipitated proteins that co-purified with SLX4 were eluted from the beads with LDS (Lithium dodecyl sulfate) sample buffer were loaded on NuPAGE 4-12% Bis-Tris acrylamide gels (Life Technologies) to stack proteins in a single band. Following staining with Imperial Blue (Thermo Fisher Scientific), protein bands were excised from the gel and gel pieces were submitted to in-gel trypsin digestion following cysteines reduction and alkylation (3). Peptides were extracted from the gel and vacuum dried. Samples were reconstituted with 0.1% trifluoroacetic acid in 4% acetonitrile and analysed by liquid chromatography (LC)-tandem mass spectrometry (MS/MS) using an Orbitrap Fusion Lumos Trisboid Mass Spectrometer (Thermo Fisher Scientific, San Jose, CA) online with a nanoRSLC Ultimate 3000 chromatography system Thermo Fisher Scientific, Sunnyvale, CA). Peptides were separated on a Thermo Scientific Acclaim PepMap RSLC C18 column (2µm, 100A, 75 µm x 50 cm). For peptide ionization in the EASY-Spray nanosource in front of the Orbitrap Fusion Lumos Trisboid Mass Spectrometer, spray voltage was set at 2.2 kV and the capillary temperature at 275 °C. The Orbitrap Lumos was used in data dependent mode to switch consistently between MS and MS/MS. Time between Masters Scans was set to 3 seconds. MS spectra were acquired with the Orbitrap in the range of m/z 400-1600 at a FWHM resolution of 120 000 measured at 400 m/z. AGC target was set at 4.0e5 with a 50 ms Maximum Injection Time. For internal mass calibration the 445.120025 ions was used as lock mass. The more abundant precursor ions were selected and collision-induced dissociation fragmentation was performed in the ion trap to have maximum sensitivity and yield a maximum amount of MS/MS data. Number of precursor ions was automatically defined along run in 3s windows using the "Inject Ions for All Available parallelizable time option" with a maximum injection time of 300 ms. The signal threshold for an MS/MS event was set to 5000 counts. Charge state screening was enabled to exclude precursors with 0 and 1 charge states. Dynamic exclusion was enabled with a repeat count of 1 and duration of 60 s.

Raw files generated from mass spectrometry analysis were processed with Proteome Discoverer 1.4.1.14 (Thermo fisher Scientific) to search against the human protein proteome of the swissprot database (20,368 entries, extracted from Uniprot on november 2019). Database search with Mascot were done using the following settings: a maximum of two trypsin miss cleavage allowed, methionine oxidation and protein N-terminus acetylation as variable modifications and cysteine carbamidomethylation as fixed modification. A peptide mass tolerance of 6 ppm and a fragment mass tolerance of 0.8 Da were allowed for search analysis. Only peptides with high stringency Mascot scores were selected for protein identification. False discovery rate was set to 1% for protein identification.

*Identification of the N-terminus of SLX4ΔNter:*

To identify the amino-terminal end of SLX4 protein, SLX4 was first immunoprecipitated from FiTo (WT SLX4) or KO30 (SLX4ΔNter) nuclear extracts. Proteins were separated on NuPAGE 4-12% Bis-Tris
acrylamide gels (Life Technologies) and following imperial blue staining, the upper part of the gel corresponding to proteins between MW 150 and 300 kDa was cut in 4 separate bands (respectively bands 1 to 4 for WT and 5 to 8 for KO30 extracts). Each band was digested as previously described and analyzed by liquid chromatography (LC)-tandem MS (MS/MS) using a Q Exactive Plus Hybrid Quadrupole-Orbitrap online with a nanoLC Ultimate 3000 chromatography system (Thermo Fisher Scientific™, San Jose, CA). 5 microliters corresponding to 33 % of digested protein were injected on the system. After pre-concentration and washing of the sample on a Acclaim PepMap 100 column (C18, 2 cm × 100 μm i.d. 100 A pore size, 5 μm particle size), peptides were separated on a LC EASY-Spray column (C18, 50 cm × 75 μm i.d., 100 A, 2 μm, 100A particle size) at a flow rate of 300 nL/min with a two steps linear gradient (2-20% acetonitrile/H2O; 0.1 % formic acid for 40 min and 20-40% acetonitrile/H2O; 0.1 % formic acid for 10 min). For peptides ionization in the EASYSpray source, spray voltage was set at 1.9 kV and the capillary temperature at 250 °C. All samples were measured in a data dependent acquisition mode. Each run was preceded by a blank MS run in order to monitor system background. The peptide masses were measured in a survey full scan (scan range 375-1500 m/z, with 70 K FWHM resolution at m/z=400, target AGC value of 3.00×10^6 and maximum injection time of 100 ms). Following the high-resolution full scan in the Orbitrap, the 10 most intense data-dependent precursor ions were successively fragmented in HCD cell and measured in Orbitrap (normalized collision energy of 27 %, activation time of 10 ms, target AGC value of 1.00×10^5, intensity threshold 1.00×10^4 maximum injection time 100 ms, isolation window 2 m/z, 17.5 K FWHM resolution, scan range 200 to 2000 m/z). Dynamic exclusion was implemented with a repeat count of 1 and exclusion duration of 10 s.

Raw files generated from mass spectrometry analysis were processed with Proteome Discoverer 1.4.1.14 (Thermo fisher Scientific) to search against the human protein proteome of the swissprot database (20,368 entries, extracted from Uniprot on november 2019) modified by the addition of 85 SLX4 sequences. The Q8IY92 uniprot entry corresponding to the entire sequence 1-1834 of the SLX4 protein was used to create and add artificial 85 different SLX4 sequences corresponding to amino-terminal truncated proteins deleted from 301 to 386 first amino-acids, each sequence differing by the incremental deletion of 1 amino-acid. First sequence named Q8IY92-302 corresponds for example to 302N-N1834 SLX4 sequence and Q8IY92-387 to 387F-N1834 SLX4 sequence. Database search with Mascot were done using the following settings: a maximum of two trypsin miss cleavage allowed, methionine oxidation and protein N-terminus acetylation as variable modifications and cysteine carbamidomethylation as fixed modification. A peptide mass tolerance of 10 ppm and a fragment mass tolerance of 0.8 Da were allowed for search analysis. Only peptides with high stringency Mascot scores were selected for protein identification. False discovery rate was set to 1% for protein identification. To compare SLX4 sequence coverage and identify N-terminal sequence for both WT and truncated form of SLX4, one search of raws corresponding to bands 1 to 4 (WT) was compared to corresponding search of raws 5 to 8 from N-terminally truncated SLX4 (KO30).
SUPPLEMENTARY FIGURE LEGENDS

Figure S1 – related to Fig 1. SLX4 interaction with MSH2. (A) Number of peptide spectrum matches (PSM) of MSH2 and MSH6 found in endogenous SLX4 immunoprecipitates using whole cell extracts from control or SLX4-depleted HeLa Fito cells. (B) Interaction of MSH2 with the N-terminus of SLX4. The scheme indicates the various FLAG-tagged fragments of SLX4 transfected in HeLa Fito cells and immunoprecipitated before WB analysis. n/a indicates a non-relevant SLX4 mutant used for another research project. (C) SLX4 UBZ domains are dispensable for MSH2 interaction. HeLa Fito cells were transiently transfected with YFP-SLX4 WT or a UBZ mutant (C296S, C299S, C336S, C339S) before GFP pull-down. The asterisk represents an aspecific band recognized by the anti-XPF antibody. (D) The lever 1 domain of MSH2 is required for SLX4 interaction. HeLa Fito cells were transfected with FLAG-SLX4 together with the indicated GFP-MSH2 constructs before GFP trap affinity purification and western blotting.

Figure S2 – related to Fig 2. CRISPR-Cas9 inactivation of SLX4 (A) Experimental procedure aiming at knocking out SLX4 in HeLa cells using CRISPR-Cas9 and homology-directed repair allowing the insertion of a Puromycin-containing plasmid at the SLX4 locus. (B) MMC sensitivity assay of selected Puromycin-resistant clones. Cells were seeded in 6 well plates, treated with MMC (5 ng/ml for 24 h) before drug wash out and addition of fresh medium. Cells were fixed at day 5. The red rectangles indicate clones displaying a severe MMC hypersensitivity. (C) and (D) Western blots of SLX4 in selected Puromycin-resistant clones using an anti-SLX4 recognizing a C-terminal epitope. The blue asterisk indicates a recurrent band with a lower molecular weight (MW) found in several clones. One clone (KO30) initially displayed an apparent knock-out of SLX4 but the same anti-SLX4 reacting band with a lower MW was found in subsequent analysis.

Figure S3 – related to Fig 2. Characterization of KO30 cells expressing SLX4ΔNter. (A) Immunoprecipitation of SLX4 and SLX4ΔNter from Fito and KO30 cells, the asterisk indicates an aspecific band that can be detected by the anti-SLX4 antibody in Fito input and migrates at the same size of SLX4ΔNter but that is not immunoprecipitated. (B) Lack of SLX4 N-terminus in KO30 cells. SLX4 immunoprecipitates from HeLa Fito or KO30 cells were analysed by mass spectrometry (MS). As indicated, peptides in red or orange found only in parental Fito cells were overrepresented in SLX4 N-terminus. (C) Scheme of the SLX4 locus in Fito or in CRISPR-Cas9 targeted KO30 cells. PCR primers used in (D) are indicated. (D) Disruption of Exon 3 integrity and plasmid insertion in KO30 cells. PCR analysis of genomic DNA from Fito and KO30 cells were performed with the indicated primers. (E) Model for the generation of SLX4ΔNter in KO30 cells. Alternative Transcription Start sites (Alt-TSS) were found by TSS-seq in several cell lines and reported in the DBTSS (DataBase of Transcriptional Start Sites: https://dbtss.hgc.jp/). Large triangles represent a strong cluster of TSS in the beginning of intron 3 while small triangles represent other TSS in the end of intron 3 and beginning of exon 4. These alt-TSS are compatible with the N-terminal proximal SLX4
peptides found in KO30 cells. (F) Candidate alternative Translation Initiation Site (alt-TIS) matching the Kozac sequence consensus. Coloured bases are compatible with a Kozac sequence consensus (1). (G) This alt-TIS used in KO30 cells would generate the indicated tryptic peptide that was however found only in FITo cells, possibly because it is post-translationally modified (see also supplementary results). (H) Peptide spectrum of the proximal peptide containing an acetylated initiator Methionine found only in KO30 cells after immunoprecipitation of endogenous SLX4 from HeLa FITo or KO30 nuclear extracts and mass spectrometry analysis.

**Figure S4 — related to Fig 3 and Fig 4.** (A) Western blot showing a representative result of MSH2 or SLX4 depletion in HeLa FITo cells transfected with the indicated siRNAs in clonogenic survival assays. (B) Clonogenic survival in response to MMC (5 ng/ml) or Melphalan (500 nM) upon MSH2 or SLX4 depletion (n=3 to 5 experiments, mean ± SD are indicated)

**Figure S5 — related to Fig 5.** SLX4 interacts with MSH2 through a SHIP box. (A) Alignment of human MSH2 and S. cerevisiae Msh2 proteins showing the conservation of M453/M470 and its surrounding residues. (B) Structure of the MSH2-MSH6 complex. The figure was generated with PyMol using PDB coordinates 2O8C. MSH2 is in lavender with its lever 1 domain in marine blue, MSH6 is in cyan and the DNA strands in yellow and orange. The side chain of MSH2 M453 residue in lever 1 is shown as pink (carbon atoms) and yellow (sulfur atom) spheres while Mg²⁺ ions are shown as grey spheres next to the ADP molecules shown in stick representation. This structure 2O8C represents “Human MutSα bound to ADP and an O6meG-T mispair” and comes from Warren JJ et al. (4). (C) Scheme of the in vitro MMR assay and example of an expected result, see Material and Methods for details. (D) In vitro mismatch repair assay using a plasmid containing a G/T mismatch and a 5’ nick incubated with nuclear extracts (NE) from HeLa FITo cells as described in material and methods. The WT SLX4 SHIP peptide was added to the reaction at the indicated concentration. DNA was purified and digested with Apal and Pvull. Repair of the mismatch restores the PvuII site and produces two bands of 1.55 kb and 1.03 kb on an agarose gel. The percentage of repair is indicated. (E) Immunoprecipitation of GFP-MSH2 overexpressed in HeLa FITo or KO30 cells. (F) Immunoprecipitation of GFP-MSH2 overexpressed in KO30 cells complemented with FHA-SLX4 WT or FHA-SLX4ΔMSH2bd. Expression of exogenous SLX4 was achieved with 10 ng/ml doxycycline.

**SUPPLEMENTARY REFERENCES**

1. Wegrzyn,J.L., Drudge,T.M., Valafar,F. and Hook,V. (2008) Bioinformatic analyses of mammalian 5’-UTR sequence properties of mRNAs predicts alternative translation initiation sites. *Bmc Bioinformatics*, 9, 232.
2. Arnesen, T., Damme, P.V., Polevoda, B., Helsens, K., Evjenth, R., Colaert, N., Varhaug, J.E., Vandekerckhove, J., Lillehaug, J.R., Sherman, F., et al. (2009) Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proc National Acad Sci*, 106, 8157–8162.

3. Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Analytical chemistry*, 68, 850–858.

4. Warren, J.J., Pohlhaus, T.J., Changela, A., Iyer, R.R., Modrich, P.L. and Beese, L.S. (2007) Structure of the Human MutSα DNA Lesion Recognition Complex. *Mol Cell*, 26, 579–592.
Figure S1

A

| SUM OF PSM number (3 runs per exp) | experiment 1 | experiment 2 |
|-----------------------------------|--------------|--------------|
| siRNA:                            | siLUC        | siLUC        |
| Ab for IP:                        | Ab1+2        | Ab1+2        |
| Q8IY92 Structure-specific endonuclease subunit SLX4 | 111          | 44           |
| Q92889 DNA repair endonuclease XPF | 65           | 11           |
| P43246 DNA mismatch repair protein MSH2 | 13           | 0            |
| P52701 DNA mismatch repair protein MSH6 | 13           | 0            |

B

C

| Input | GFP Pull-down |
|-------|---------------|
| YFP-SLX4 | YFP-SLX4 |

Transfection:

- αMSH2
- αXPF
- αSLX4
- αFLAG
Figure S1

Transfection: (+FLAG-SLX4)

| Input | GFP-MSH2 | GFP pull-down |
|-------|----------|---------------|
|       |          |               |

- **αSLX4**
  - Input: GFP FL 1-620 1-550 1-460 1-310
  - GFP pull-down: GFP FL 1-620 1-550 1-460 1-310

- **αMSH6**
  - Input: GFP FL 1-620 1-550 1-460 1-310
  - GFP pull-down: GFP FL 1-620 1-550 1-460 1-310

- **αMSH3**
  - Input: GFP FL 1-620 1-550 1-460 1-310
  - GFP pull-down: GFP FL 1-620 1-550 1-460 1-310

- **αGFP**
  - Input: GFP FL 1-620 1-550 1-460 1-310
  - GFP pull-down: GFP FL 1-620 1-550 1-460 1-310

KDa:
- ~205
- ~120
- ~85
- ~65
- ~50
- ~30
Figure S2

A

HeLa Flp-In T-Rex (FITo) \[\rightarrow\] HeLa Flp-In T-Rex (FITo) Transfection CRISPR-Cas9 and Homology-Directed Repair (HDR) plasmids \[\rightarrow\] Selection Individuals clones selected with Puromycin \[\rightarrow\] Screening Western blot MMC sensitivity

B

FITo (parental)

Untreated MMC

#32

#12

#40

#1

C

| Clones | FITo 1 2 9 10 12 13 32 38 40 | αSLX4 | αSLX4 (low exp) |
|--------|-----------------------------|--------|-----------------|
| MMC    | 1 2 9 10 12 13 32 38 40     | 250 kDa| 250 kDa         |

D

| Clones | FITo 4 5 13 18 22 27 30 | αSLX4 | αXPF | Ponceau |
|--------|-------------------------|--------|------|---------|
| MMC    | 4 5 13 18 22 27 30      | 250 kDa|      |         |
Figure S3

A

|          | Input | IP IgG | IP SLX4 |
|----------|-------|--------|---------|
| HeLa:    | FIT0  | KO30   |         |
| αSLX4   | *     |        |         |
| αXPF    |       |        |         |

**Figure S3**

**A**

| HeLa: | **FIT0** | **KO30** |
|-------|----------|----------|
| αSLX4 | *        |          |
| αXPF  |          |          |

**B**

**RED/ORANGE:** peptides found only in HeLa FIT0 (parental) (in bold: peptides found at least 4 times)

**GREEN:** peptides found in both HeLa FIT0 and KO30

**BLUE:** peptides found only in KO30

>sp|081Y92|SLX4_HUMAN
MKLsvNCAQLGFYGLSLSLSACPGIDPSSDQPSLKLGTGQMDDESDFKELCASFFQ
RVKKHGKEVSGERTKQASANCTIQKSLKRTQATKTKTLQGPAEKKPPGSQAPRT
KQRVTKWQAEPASHVNEGGLASADPPQVLRETAQNTQTNQQREFSPNLSREKTRE
VNSDQPPSCCLTAVPSKPRATAQLVLQRMQKFPRADPRILRHAESCSELAAREEN
VKPDQEEMAGNLYGQLPAPQDGSVDAVALTLQGQFARVGASAHDDSELEEGLFFCCICQ
KLNSAMNVRQHVRNLQCDEAEKRTPLQVQIEPCICGKFPFLTKLRRTHLQKCAVKM
EVGQPLLQQAVRLQTAQPEGSSPSNFSFDHRSGLKRRGPSKKEPRRRKVDDEAPSD
LLVAMALRSEMPEGAAPVLRLESASERTRPEAEKNSRRKKPPVSPPPLLLVQSDTG
RQI3EDRVALLSEEFVSSTPPLASLRKENGTERAQCPETPRKQSLWEGSAEGTA
AMEDFTRYALVPPLVPQPAQGLPEPVLVPLPQHPHSELSRSPHHTPAGCSCRG
SPASASQREHQALQIDLVDLAREGQASASPDGSQGGLAGSEGTCAGDLVPPGLPTGFVSQ
DKHPDGRGRTLSLSLGLVADFAGMVMNPHELSDQFQTDSEVLYAHKVFVLYARCLPQY
VMNGFSAVEDGVTQRLGDVSTEAAARTPHHLATFDGLPGSSLSLLSHLRFGVS
ELVHLCEQPFIASTDEGKIFEEKEAEENGESRAENFQELLRSMNWADDEEAETLKLKSDHE
EDQENVNEEMEEEYEAFTAQKLLQERAAAGADEADWLEGGPSVPSQQLAVGQQKQN
DKVEEMPLEEPGRDEAAATTMKEQGQCALPPQPGQHSARGAEAPAEEALGHSSCSS
PSRDCQAERKEGSLPHSDADGYEQLFSSTQGEISEPSQTSEPEEQSGAVRERGLEVSH
RLAPWQAPSHPCRFLLGPPQGSPRGSHTSGSGLSTLSTRPRGTSQVSPLLLSAPVPS
KQKRDSR3ILTLSEPHQKKGKERSVLECRNKGVLMPFESPSIDLTQSNPDSSSRSSQK
SSKTNNEDEVLLSDDELEETKMKISSDPLEEKKALEISPRSCFELSIIVDAD
QEPSQSPPRSEAVQLOQEDGALPENRGSLRGAPWLCDERESSPSSEASTDTTDSWLVPAT
PLASRSRDCCSSQTQISSLRSLGALQVHTPRASVNVREGNEVAQKFSRSVIRQTPFQQT
PPSSCLTVPSTGDSRQHGRSPSRPPIPQGHPSPLAPHISGDRAHRSRFLHKFSPPG
PSFLAQTPQAGEVEVGDSDEQEVASHQANSRPLDSPDPIPIDDCCWHMEIPSSPIDH
WNLERTGLPLTSSPSRNNMAADRSRDCR5PGLLTTPRQSCSTTQRKLQKESSSGASLGN
SRPSFLNSALWVMGDGEQRPETPFPAPMQPSAGQAQPKPELETTFPGANKRNKKLPVKPI
TPMPQYSIMETPVKELDRGFVRPLKPRQMLKLKEIFYQHYTQLTSLDEDESSQSSQPL
LQAPHCQQLTQKPSRAGVHQAETGPGAhRPGPAKTGKPRHRKHIESITTSPSR
SPTKEAPPRLNDNADQPIQASQESVSATSVDGSSLSLSQSSQSSCSFGAFASEAGEEIEG
SASQQAVQAADTEALRCRY INCLUDED PALYQKVLQQYPFELRELQAELEHRQGLVSSRPLLDF
LDTHCITFTTAATRREKLLQGRRQPRGKXVER
**Figure S3**

C

WT

KO30

D

Oligos: Fwd+Rev

Genomic DNA:

Expected size of PCR products:

KO30 SLX4 mRNA

E

Looking for nearby potential Kozak sequence

KO30 SLX4 protein start?

SLX4ΔNter
Figure S3

F

altTIS:   CTGTGAAG ATG GAGG TTGGCCCC
         M E V G P
         (360)

G

...NLSAMNVTRREQHVRCLDEAEKTLRPSVPQIPECPICGKPFLTLKSRTSHLKQCAVK
MEVGPQLLLQAVRLQTAQPEGSSSPMFSFDHSGLKRGPSTKKEPRKRKVDEAP
SEDLLVAMALSR...

        ↓
        MEVGPQLLLQAVR: n=14 in FItO
        never found in KO30

        ↓

Post translational modification?
Nter acetylation?

H

Peptide spectrum of
Ac-MEVGPQLLLQAVR
found exclusively in KO30 cells

|          | number of PSM |
|----------|---------------|
|          | MEVGPQLLLQAVR | Ac-MEVGPQLLLQAVR |
| FItO     | 6             | 0                 |
| KO30     | 0             | 3                 |
Figure S4

A

B

% survival

MMC Melphalan

siLUC siMSH2 siSLX4-3
Figure S5

A

| Human | 407 | LYGINQLPNVQAA---LEKHEGKH--OXLALLAVFVTPLTLRSDSFSPFQEMIETLDM | 460 |
|-------|-----|-------------------------------------------------------------|-----|
| Yeast | 418 | IYQFSKRIPEIVQUFTSPLEDSPTEPVNEELVRSVWLAPLSHVEPLSKFEEVETTVDL | 477 |

B

C

| PvuII | G | + HeLa nuclear extracts (50 µg + dNTP, ATP...) +/- SLX4 peptide |
|-------|---|-----------------------------------------------------------------|
| 5' nick (132bp) | T | MMR (5' directed) |
| ApaLI | |

Nuclear extracts: - +

Linear \(\rightarrow\) (ApaLI) 2579bp 1549bp 1030bp

\(\leftarrow\) unrepaired \(\leftarrow\) Repair products (PvuII&ApaLI)

DNA precipitation/DNA Digestion
Figure S5

D

HeLa NE
+ SLX4 WT SHIP

(kb)

150
225
150
225
102

\( \alpha \) EXO1

Input

GFP-
MSH2

GFP pull-down

GFP-
MSH2

KO30+FHA-SLX4:

\( \alpha \) GFP

\( \alpha \) SLX4

FL \( \rightarrow \) \( \Delta \) Nter

\( \alpha \) GF

E

F

KO30+FHA-SLX4:

\( \alpha \) EXO1

\( \alpha \) SLX4

(FHA-SLX4)

\( \alpha \) GFP

(GFP-MSH2)

% repair: 22 25 3 <1 <1

HeLa NE
+ SLX4 WT SHIP

unrepaired

repair

products

(kb)

1
1.5
2
2.5
3

0.8

102