DNA interstrand cross-link repair requires replication-fork convergence

Jieqiong Zhang[1], James M Dewar[1], Magda Budzowska[1], Anna Motnenko[2], Martin A Cohn[2] & Johannes C Walter[3]

DNA interstrand cross-links (ICLs) prevent strand separation during DNA replication and transcription and therefore are extremely cytotoxic. In metazoans, a major pathway of ICL repair is coupled to DNA replication, and it requires the Fanconi anemia pathway. In most current models, collision of a single DNA replication fork with an ICL is sufficient to initiate repair. In contrast, we show here that in *Xenopus* egg extracts two DNA replication forks must converge on an ICL to trigger repair. When only one fork reaches the ICL, the replicative CMG helicase fails to unload from the stalled fork, and repair is blocked. Arrival of a second fork, even when substantially delayed, rescues repair. We conclude that ICL repair requires a replication-induced X-shaped DNA structure surrounding the lesion, and we speculate on how this requirement helps maintain genomic stability in S phase.

DNA ICLs involve a covalent linkage between the Watson and Crick strands of DNA. When left unrepaired, a small number of ICLs can kill a mammalian cell[1]. This cytotoxicity is widely exploited for cancer chemotherapy, which uses bifunctional cross-linking agents such as the nitrogen mustards, platinum compounds and mitomycin C[1]. It has been proposed that endogenous metabolites such as reactive aldehydes also cause ICLs *in vivo*[2]. Failure to repair ICLs and other lesions might be the underlying cause of Fanconi anemia, a rare syndrome marked by bone-marrow failure and cancer predisposition[3].

In vertebrate cells, a major pathway of ICL repair occurs in the S phase of the cell cycle[4]. This mechanism requires the Fanconi anemia pathway, structure-specific endonucleases, translesion DNA polymerases and recombinases[2,5]. We previously showed that in *Xenopus* egg extracts a plasmid containing a site-specific cisplatin ICL (pICL) undergoes replication-coupled ICL repair (**Supplementary Fig. 1a** and ref. 6). In this system, replication initiates in a sequence-nonspecific manner, and two replisomes converge on the ICL. Their leading strands stall ~20–40 nt from the lesion ('–20' position), owing to steric hindrance by the CDC45, MCM2–7, GINS (CMG) helicase, which translocates on the leading-strand template ahead of DNA polymerase[7]. The BRCA1–BARD1 complex then promotes the dissociation of CMG from the stalled forks[8], and this is followed by leading-strand extension to within 1 nt of the ICL ('approach' to '–1' position). Next, ubiquitinated FANCI–FANCD2 binds chromatin and helps recruit the XPF–ERCC1–SLX4 complex[9,10]. XPF–ERCC1 incises one parental strand ('unhooking') and allows lesion bypass on the other parental strand by translesion DNA polymerases. Finally, the double-stranded DNA break generated by incisions is repaired by homologous recombination[11]. In this cell-free system, the cisplatin adduct remains attached to one parental strand.

The use of a small plasmid to model ICL repair in egg extracts inevitably leads to rapid convergence of two DNA replication forks on the lesion. In contrast, *in vivo,* where the average interorigin distance is large (~100 kb)[12], one replication fork should generally encounter an ICL well before a second fork arrives. Therefore, although convergent forks are generally viewed as being able to trigger ICL repair, it is widely assumed that a single fork is also sufficient[2,4,5,13–18] (**Supplementary Fig. 1b**). In apparent agreement with single fork–induced repair, a psoralen ICL flanked on one side by a replication roadblock (EBNA1 protein bound to FR repeats) can be repaired[19]. However, this result is ambiguous because fork arrest by EBNA1 is incomplete[20]. In addition, cell-free replication of a psoralen-ICL plasmid has suggested that a single fork can trigger incisions, but these were not shown to require the Fanconi anemia pathway or to promote repair[21].

Here, we set out to examine what happens when only a single fork strikes an ICL. We found that in *Xenopus* egg extracts, one DNA replication fork is completely inert for ICL repair. Specifically, the CMG complex is not unloaded from a single fork stalled at an ICL, the leading strand fails to approach to the lesion, and no downstream repair events are detected. Importantly, ICL repair is still productive when there is a major delay between the arrival of the first and second forks, as would occur *in vivo*. Finally, we show that the order in which two forks strike the ICL does not affect the mechanism of repair. Together with previous results[22], our data indicate that formation of an X-shaped structure surrounding an ICL is the essential trigger for ICL repair.

RESULTS

A single fork stalled at an ICL fails to undergo approach

We wanted to directly compare what happens when one or two replication forks collide with an ICL. To this end, we constructed pICL-*lacO*, in which we placed an array of 48 *lac* operator (*lacO*) sites

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, USA. [2]Department of Biochemistry, University of Oxford, Oxford, UK. [3]Howard Hughes Medical Institute, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, USA. Correspondence should be addressed to J.W. (johannes_walter@hms.harvard.edu).
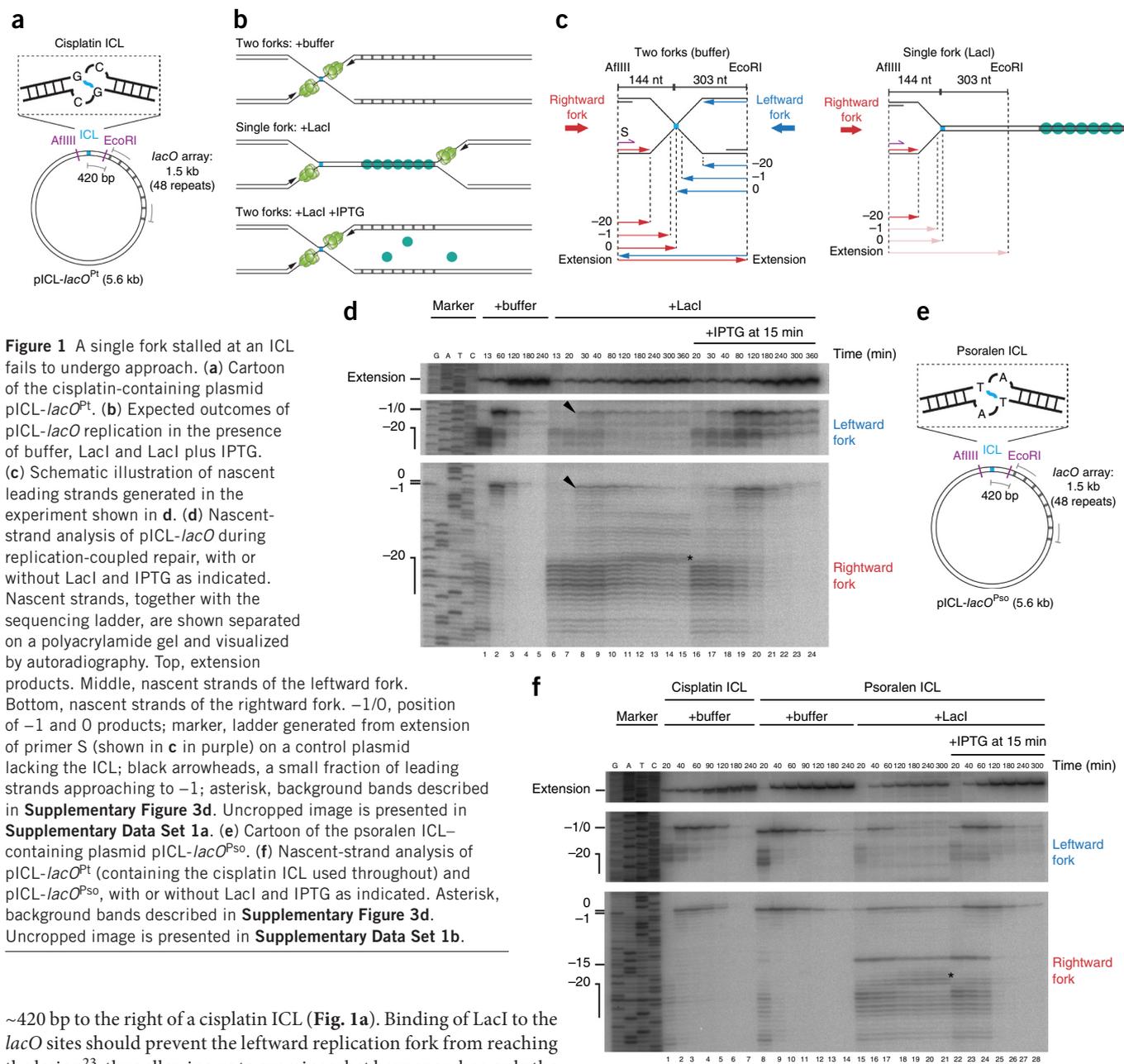
Figure 1 A single fork stalled at an ICL fails to undergo approach. (**a**) Cartoon of the cisplatin-containing plasmid pICL-*lacO*Pt. (**b**) Expected outcomes of pICL-*lacO* replication in the presence of buffer, LacI and LacI plus IPTG. (**c**) Schematic illustration of nascent leading strands generated in the experiment shown in **d**. (**d**) Nascent-strand analysis of pICL-*lacO* during replication-coupled repair, with or without LacI and IPTG as indicated. Nascent strands, together with the sequencing ladder, are shown separated on a polyacrylamide gel and visualized by autoradiography. Top, extension products. Middle, nascent strands of the leftward fork. Bottom, nascent strands of the rightward fork. −1/0, position of −1 and 0 products; marker, ladder generated from extension of primer S (shown in **c** in purple) on a control plasmid lacking the ICL; black arrowheads, a small fraction of leading strands approaching to −1; asterisk, background bands described in **Supplementary Figure 3d**. Uncropped image is presented in **Supplementary Data Set 1a**. (**e**) Cartoon of the psoralen ICL–containing plasmid pICL-*lacO*Pso. (**f**) Nascent-strand analysis of pICL-*lacO*Pt (containing the cisplatin ICL used throughout) and pICL-*lacO*Pso, with or without LacI and IPTG as indicated. Asterisk, background bands described in **Supplementary Figure 3d**. Uncropped image is presented in **Supplementary Data Set 1b**.

~420 bp to the right of a cisplatin ICL (**Fig. 1a**). Binding of LacI to the *lacO* sites should prevent the leftward replication fork from reaching the lesion[23], thus allowing us to examine what happens when only the rightward fork encounters the ICL (**Fig. 1b**, middle). In the absence of an ICL, the LacI array inhibited replication-fork progression for at least 3 h (**Supplementary Fig. 2a** and **Supplementary Fig. 2b**, lanes 5–10)[23]. When we added IPTG 15 or 75 min after the initiation of DNA replication, the stalled replication forks restarted and completed DNA synthesis (**Supplementary Fig. 2b**, lanes 11–17 and lanes 18–24, respectively), although the rate of replication was reduced, probably because of residual binding of LacI to DNA in the presence of IPTG. Therefore, the LacI array can be used to control access of the leftward fork to the lesion in pICL-*lacO*.

We first investigated whether a single fork could undergo approach, an early event in ICL repair (**Supplementary Fig. 1a**). We preincubated pICL-*lacO* with buffer or LacI and then allowed replication to proceed in egg extract containing [α-32P]dATP to radiolabel nascent DNA strands. Then we digested replication intermediates with AflIII and EcoRI and separated them on a denaturing polyacrylamide gel to monitor nascent-strand synthesis at nucleotide resolution (**Fig. 1c**).

Without LacI, both the leftward and rightward leading strands reached the −20 position, after which they approached to the −1 position before undergoing extension past the lesion (**Fig. 1d**, lanes 1–5)[6]. As expected, in the presence of LacI, arrival of the leftward leading strands was strongly reduced (by ~74% on average) (**Fig. 1d**, middle, comparison of lanes 1 and 6). Strikingly, although LacI did not affect arrival of rightward leading strands at the −20 position (**Fig. 1d**, bottom, comparison of lanes 1 and 6), their approach to −1 was dramatically inhibited (**Fig. 1d**, bottom, lanes 6–15). Specifically, ~70% of rightward strands failed to approach by 120 min, and more than 50% remained stalled at −20 for 6 h (**Supplementary Fig. 3a**). In agreement with this inhibition of approach, there was on average a 70% reduction in extension products (**Fig. 1d**, top, lanes 6–15). Although a fraction of leftward and rightward leading strands did approach to −1 (**Fig. 1d** and **Supplementary Fig. 3a**), this could be explained by the arrival of 26% of leftward forks at the lesion (**Fig. 1d**, middle, lane 6), thus resulting in fork convergence. We speculate that residual arrival of

Figure 2 CMG is not evicted from a single fork stalled at an ICL. (**a**) Schematic of primers used in ChIP. (**b**) MCM7 and CDC45 ChIP analysis at different time points during repair. pICL-*lacO* was replicated with or without LacI, and IPTG was added immediately before the 20-min time point, as indicated (green arrows). A repetition of this experiment is shown in **Supplementary Figure 5a**. Ctrl, control; IP, immunoprecipitation.
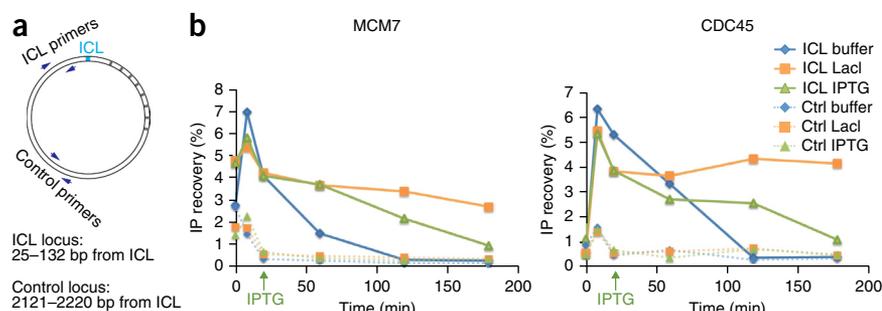
leftward forks is due to occasional origin firing between the ICL and the LacI array or to incomplete inhibition of leftward-fork progression by the LacI array, especially at late time points (**Supplementary Fig. 2b**, lanes 9 and 10). Importantly, when we added LacI immediately after most forks had converged, leading strands underwent normal approach and extension (**Supplementary Fig. 3b,c**). Therefore, once forks have converged at the ICL, the LacI array does not inhibit repair. When we mixed a plasmid containing a *lacO* array (p*lacO*) with pICL in the presence of LacI, approach on pICL was unaffected, thus demonstrating that the LacI array does not inhibit approach in *trans* (**Supplementary Fig. 3d**). Approach was also inhibited when a single fork encountered a psoralen ICL (**Fig. 1e,f**). Our results show that in *Xenopus* egg extracts, a single fork stalled at an ICL is not able to undergo approach, the first event of ICL repair.

A single fork stalled at an ICL remains competent for repair

In vivo, one fork will usually strike an ICL well before a second fork arrives. To mimic this situation *in vitro*, we allowed one fork to strike an ICL in the presence of LacI. We then disassembled the LacI barrier at different times by adding IPTG and then measured approach. When we added IPTG 15 or 75 min after replication initiation, approach and extension were restored (**Fig. 1d,f** and **Supplementary Fig. 4**). Thus, a single stalled fork remains competent for ICL repair for extended periods of time. We conclude that ICL repair is productive even when there is an extensive delay between the arrival of the first and second forks, as would normally occur *in vivo*.

CMG is not evicted from a single fork stalled at an ICL

We previously showed that leading-strand approach requires dissociation of CMG from the stalled replisomes[7,8]. We therefore postulated that the failure of a single fork to undergo approach might be caused by defective CMG dissociation. To test this idea, we examined CMG localization at the ICL locus and at a control locus distal to the ICL, by using chromatin immunoprecipitation (ChIP) (**Fig. 2**). In the presence of buffer, the MCM7 and CDC45 subunits of the CMG helicase accumulated at the ICL and then dissociated (**Fig. 2b** and **Supplementary Fig. 5a**, solid blue lines)[7]. In the presence of LacI, only ~30% of MCM7 and CDC45 dissociated from the ICL (in agreement with residual fork convergence), whereas the majority persisted (**Fig. 2b** and **Supplementary Fig. 5a**, solid orange lines) and dissociated only upon IPTG addition (**Fig. 2b** and **Supplementary Fig. 5a**, solid green lines). Dissociation of CMG from the control locus was unaffected by LacI (**Fig. 2b** and **Supplementary Fig. 5a**, dashed lines). Our data indicate that CMG does not dissociate from a single fork that has stalled at an ICL, thus accounting for the failure in approach.

A single stalled fork is not incised

We next addressed whether a single fork could trigger ICL unhooking, the signature event of ICL repair. To address this question, we replicated

pICL-*lacO* with or without LacI in extract containing [α-^{32}P]dATP, separated replication intermediates on a native agarose gel and visualized them by autoradiography. Without LacI, replication-fork convergence gave rise to a discrete 'figure-eight' structure (**Fig. 3a,b**, blue arrowhead) that later disappeared as a result of FANCI–FANCD2–dependent incisions and ultimately accumulated as supercoiled plasmid (**Fig. 3b**, lanes 1–6, and refs. 6,9). In the presence of LacI, we observed the expected 'theta' intermediate due to fork stalling (**Fig. 3a,b**, green arrowhead). However, the theta structure persisted for at least 3 h (**Fig. 3b**, lanes 7–12). Replication of pQuant, a plasmid lacking an ICL and *lacO* sites, was not affected by LacI (**Fig. 3b**). These results indicate that a single fork is not subject to incisions.

To measure more directly whether parental strands were incised, we linearized replication intermediates of pICL-*lacO* with BlpI, separated them on an alkaline gel and probed by Southern blotting to visualize parental strands. In the absence of incisions, a large, X-shaped molecule should be visible (**Fig. 3c**, top, blue strands), whereas after dual incisions in one parental strand, the X-shaped species should be converted to linear forms (**Fig. 3c**, bottom, blue strand). Without LacI, the X-shaped structure declined, and linear molecules appeared (**Fig. 3d**, lanes 2–5), as expected according to previous work[9]. In contrast, in the presence of LacI, the reduction of X-shaped molecules was greatly inhibited (by ~70%), and the accumulation of linear molecules was attenuated (**Fig. 3d**, lanes 6–12, and **Supplementary Fig. 5d**). Incisions were restored by the addition of IPTG (**Fig. 3d**, lanes 13–19). We conclude that a single fork stalled at an ICL is inefficiently incised.

Next, we addressed what leads to the observed incision defect. Ubiquitinated FANCI–FANCD2 binds to chromatin and promotes recruitment of the XPF–ERCC1–SLX4 complex, which unhooks the ICL[9,10]. LacI moderately reduced FANCD2 binding to the ICL, as measured by ChIP (**Supplementary Fig. 5b**). In contrast, LacI caused a more substantial reduction in the recruitment of XPF and SLX4 to the ICL, and this effect was reversed by IPTG (**Fig. 3e** and **Supplementary Fig. 5c**). Some of the recruitment of SLX4 and XPF in the presence of LacI might be due to residual fork convergence in this condition. These results suggest that although a single stalled fork containing CMG can recruit FANCI–FANCD2, fork convergence is necessary for efficient recruitment of XPF–ERCC1–SLX4 because binding requires either CMG dissociation or the presence of an X-shaped structure.

The order of replisome arrival does not affect repair

The convergence of two forks on an ICL creates an apparently symmetrical structure (**Supplementary Fig. 1a**). However, one parental strand subsequently undergoes incision while the other acts as the template for lesion bypass (**Supplementary Fig. 1a** and ref. 6). We asked whether the order in which two forks arrive at an ICL dictates which parental strand is incised and which is used for lesion bypass. To address this question, we stalled the leftward fork with LacI for 15 or 75 min
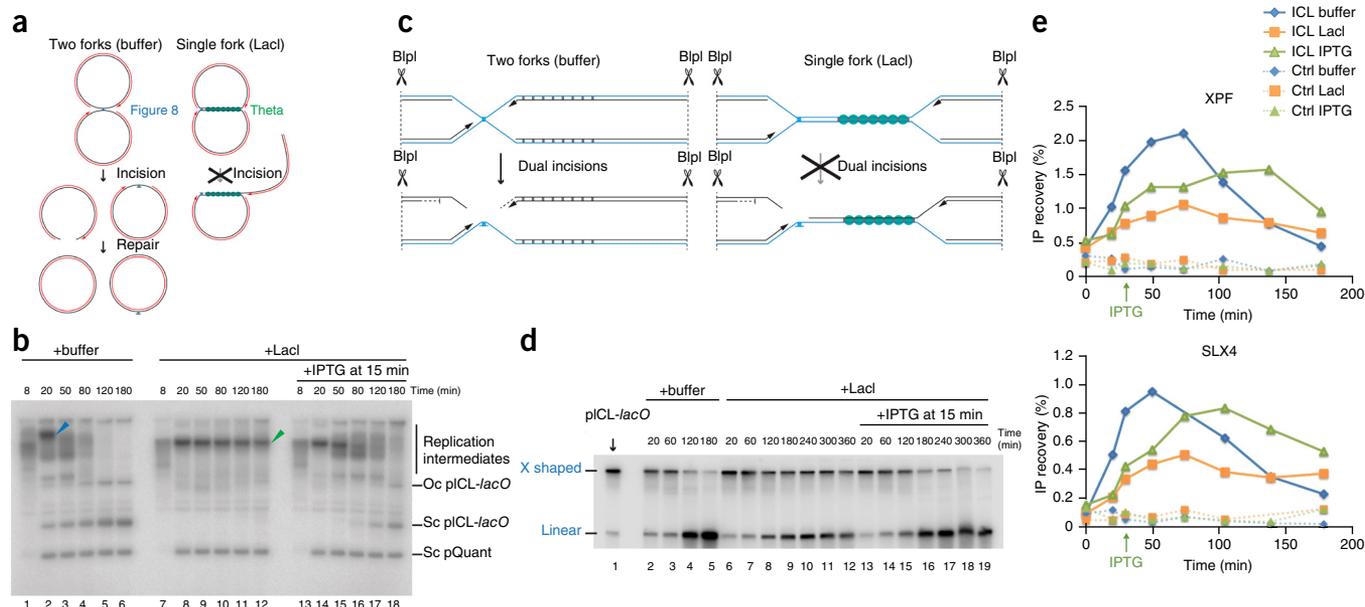
Figure 3 A single stalled fork does not undergo incisions. (**a**) Schematic of repair intermediates expected in **b**. (**b**) Replication intermediates of pICL-*lacO* separated on a native agarose gel. pICL-*lacO* and an internal control plasmid lacking the *lacO* array (pQuant) were preincubated with buffer or LacI and replicated in the presence of [α-32P]dATP. IPTG was added as indicated. Sc, supercoiled; oc, open circular; blue arrowhead, figure-eight DNA structure; green arrowhead, theta DNA structure. (**c**) Schematic of incision assay. Before dual incisions, single cutting with BlpI yields X-shaped products (blue strands, top), whereas after dual incisions, BlpI digestion yields linear molecules (blue strands, bottom). (**d**) Incision assay. pICL-*lacO* was replicated with or without LacI and IPTG, as indicated. The repair intermediates, digested with BlpI and separated on an alkaline (denaturing) gel, are visualized by Southern blotting to detect parental strands. Unreplicated pICL-*lacO* was used to generate size markers for the X-shaped structure and linear structure, which came from a small fraction of un-cross-linked background plasmids in the pICL-*lacO* preparations. Uncropped image is presented in **Supplementary Data Set 1c**. (**e**) XPF and SLX4 ChIP analysis. pICL-*lacO* and pQuant were replicated with or without LacI, and IPTG was added immediately before the 20-min time point where indicated (green arrow). At different times, samples were withdrawn for XPF and SLX4 ChIP with primer pairs for the ICL locus (**Fig. 2a**) or pQuant (Ctrl). A repetition of this experiment is shown in **Supplementary Figure 5c**.

and then released it with IPTG, so that the rightward fork reached the ICL first (**Fig. 4a**). If the rightward leading strand is then used exclusively for lesion bypass, the cisplatin adduct that persists after ICL repair[6] should be attached primarily to the bottom strand (**Fig. 4a**, rightward lesion bypass); if the two leading strands are still used for bypass with equal probability, the adduct should be detected equally on both parental strands (**Fig. 4a**, rightward and leftward lesion bypass). We digested the final repair products with AflIII and AseI, such that the top and bottom parental strands differed in size by 2 nt (**Fig. 4b**). We then separated the DNA on a sequencing gel and visualized the top or bottom parental strands by strand-specific Southern blotting. Importantly, when arrival of the leftward fork was delayed with LacI, both the bottom (**Fig. 4c**, lanes 4 and 5) and top (**Fig. 4d**, lanes 4 and 5) parental strands retained adducts. Furthermore, the ratio between adducted and unadducted strands was unaffected by LacI (**Fig. 4c,d**, comparison of lane 3 with lanes 4 and 5). We conclude that the order in which the two replisomes arrive at the ICL does not affect which parental strand is incised and which is used as the template for lesion bypass. This result suggests that a single fork remains wholly uncommitted to ICL repair until a second fork arrives. How the strand used for lesion bypass is ultimately chosen remains unclear.

DISCUSSION

In this paper, we addressed whether one or two DNA replication forks are necessary to trigger ICL repair. Because of the challenge of engineering site-specific ICLs on mammalian chromosomes and the difficulty of manipulating the abundance of DNA replication forks in cells, this question has not been addressed *in vivo*. Instead, we used a

replication-fork barrier to control the access of DNA replication forks to an ICL in *Xenopus* egg extracts, which recapitulate physiological ICL repair[2,5,6,9–11,24]. We found that, in contrast to most current models, two DNA replication forks must converge on an ICL to trigger repair. Strikingly, a single fork does not support even the first step in repair, CMG dissociation, which is presumably critical to initiate lesion bypass and to expose the ICL to the incision machinery.

Implications for ICL repair in cells

Our results raise the question of how two forks arrive at an ICL *in vivo*. Given an average interorigin distance of 100 kb (ref. 12), coordinated firing of adjacent origins[25] and an average fork rate of 1.5 kb/min (ref. 26), the maximum time delay between the arrival of the first and second forks at most ICLs should be ~60 min *in vivo*. We have shown that a single fork stalled at an ICL does not collapse and remains competent for repair for at least 60 min. Therefore, *in vivo*, ICL repair could rely on the convergence of forks from adjacent origins. In some cases, a second fork might be delivered more rapidly, owing to firing of a nearby dormant origin[27]. Interestingly, cisplatin treatment causes selective loss of telomeres[28]. This observation is consistent with our model because fork convergence cannot occur when an ICL is located beyond the last origin of replication at the chromosome end.

Recently, the Seidman group reported that a single fork can bypass a psoralen ICL without repairing it, in a manner dependent on the DNA translocase FANCM ('traverse')[22]. This observation implies that during traverse, CMG or another DNA helicase bypasses ICLs to allow fork progression. We have not observed traverse in *Xenopus* egg extracts, even on a psoralen ICL (discussion in **Supplementary Fig. 6**).
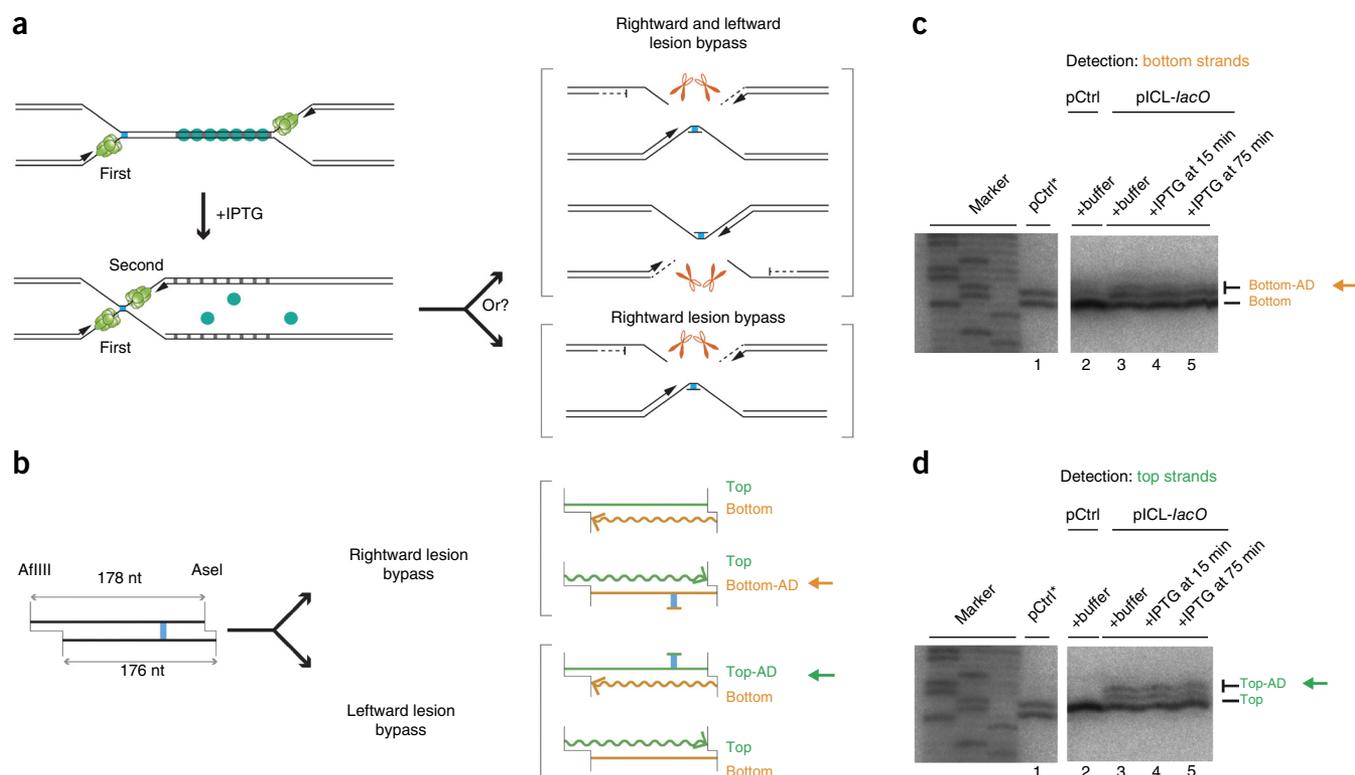
Figure 4 The order of replisome arrival at an ICL does not determine which leading strand undergoes lesion bypass. (**a**) Scheme to determine whether the order in which the two forks arrive at an ICL dictates which parental strand is used as the lesion-bypass template. (**b**) Schematic depiction of final repair products after AflIII and AseI digestion, depending on which leading strand undergoes lesion bypass. AflIII and AseI generate different-sized overhangs, thus allowing us to differentiate top (178 nt) and bottom (176 nt) strands. Top-AD or bottom-AD, top or bottom strand containing an adduct. (**c**,**d**) Strand-specific Southern blotting to detect the adducts. pCtrl or pICL-*lacO* was replicated in the presence of buffer or LacI, and IPTG was added at the indicated times. After 6 h, repair products were digested with AflIII and AseI, separated on a sequencing gel and analyzed with strand-specific Southern blotting as shown, to visualize the bottom (**c**) or top (**d**) strands. To generate size markers for the top (178 nt) and bottom (176 nt) strands (lane 1), pCtrl was replicated in the presence of [α-^{32}P]dATP (pCtrl*) and analyzed on the same sequencing gel as the strand-specific Southern blot after AflIII and AseI digestion. The absence of top (**c**, lane 2) or bottom strands (**d**, lane 2) in Southern blotting of pCtrl indicates the strand specificity of the blotting protocol. Uncropped images are presented in **Supplementary Data Set 1d**.

This failure is not due to a lack of FANCM, which is present in egg extracts[29], but perhaps it occurs because some other activity is absent in early embryos. Importantly, the lagging strand of a traversed fork is equivalent to the leading strand of a converging fork (**Supplementary Fig. 6**, comparison of purple and green strands). Thus, both traverse and fork convergence generate an X-shaped DNA structure surrounding the ICL, which we propose is the essential trigger for ICL repair. The requirement for this structure helps explain why both 5′- and 3′-directed flap endonucleases have been implicated in ICL repair[4,24].

Relationship of FANCI–FANCD2 recruitment and CMG unloading
An important question concerns the interdependence of CMG unloading with other early events in ICL repair. We recently showed that BRCA1–BARD1 is required for CMG unloading and FANCD2 recruitment at ICLs[8], thus raising the possibility that CMG eviction might be needed to make room for FANCI–FANCD2 near the lesion. Disfavoring this idea, we also showed that inhibition of approach with aphidicolin impairs CMG unloading while having little or no effect on recruitment of BRCA1–BARD1 or FANCD2 (ref. 8), thus suggesting that CMG removal is not required for efficient FANCD2 loading. This conclusion is further supported by our present finding that FANCD2 is recruited to single forks that retain CMG. Although we cannot rule out that CMG obstructs FANCI–FANCD2 binding to the ICL itself, our results suggest that it does not prevent FANCI–FANCD2 recruitment in the general vicinity of the lesion.

Implications for genome stability
What is the advantage of coupling ICL repair to fork convergence? During an unperturbed S phase, DNA replication forks are expected to stall transiently at DNA sequences and chromatin structures that are difficult to replicate[30,31]. If CMG unloading were possible from single forks, the helicase might sometimes be unloaded from transiently stalled replisomes. Given that there is no known pathway to reload the CMG complex in S phase of metazoans, the inadvertent dissociation of CMG is predicted to cause fork collapse, incomplete DNA synthesis and genome instability. Making CMG unloading absolutely dependent on fork convergence avoids this problem because the helicase would be lost only when replication is locally completed. Fork traverse past an ICL also would avoid this problem because CMG or another DNA helicase will continue unwinding on the other side of the ICL to allow completion of DNA synthesis. We recently showed that a single fork is sufficient to trigger repair of a DNA-protein cross-link (DPC), indicating that some helicase-blocking lesions can be repaired in the absence of fork convergence and CMG unloading[23]. In this case, the DPC is highly exposed and is thus amenable to destruction even in the presence of CMG, allowing fork bypass. In conclusion, our work suggests that the formation of an X-shaped

structure surrounding an ICL, either by fork convergence or fork traverse, is essential to initiate ICL repair. The unexpected failure of single stalled forks to trigger ICL repair is probably essential to avoid inadvertent fork collapse in S phase.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

AUTHOR CONTRIBUTIONS

J.M.D. generated the *lacO* array (48 *lacO* repeats) and validated its use as a replication-fork barrier; M.B. and J.Z. prepared pICL-*lacO*[Pt]; A.M. and M.A.C. prepared psoralen-cross-linked oligonucleotides; J.Z. and J.C.W. designed and analyzed the experiments; J.Z. performed all the experiments; J.Z. and J.C.W. prepared the manuscript.

1. Lawley, P.D. & Phillips, D.H. DNA adducts from chemotherapeutic agents. *Mutat. Res.* **355**, 13–40 (1996).
2. Clauson, C., Schärer, O.D. & Niedernhofer, L. Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. *Cold Spring Harb. Perspect. Med.* **3**, a012732 (2013).
3. Garaycoechea, J.I. & Patel, K.J. Why does the bone marrow fail in Fanconi anemia? *Blood* **123**, 26–34 (2014).
4. Zhang, J. & Walter, J.C. Mechanism and regulation of incisions during DNA interstrand cross-link repair. *DNA Repair (Amst.)* **19**, 135–142 (2014).
5. Deans, A.J. & West, S.C. DNA interstrand crosslink repair and cancer. *Nat. Rev. Cancer* **11**, 467–480 (2011).
6. Räschle, M. *et al.* Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* **134**, 969–980 (2008).
7. Fu, Y.V. *et al.* Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell* **146**, 931–941 (2011).
8. Long, D.T., Joukov, V., Budzowska, M. & Walter, J.C. BRCA1 promotes unloading of the CMG helicase from a stalled DNA replication fork. *Mol. Cell* **56**, 174–185 (2014).
9. Knipscheer, P. *et al.* The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* **326**, 1698–1701 (2009).
10. Klein Douwel, D. *et al.* XPF-ERCC1 acts in unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. *Mol. Cell* **54**, 460–471 (2014).
11. Long, D.T., Räschle, M., Joukov, V. & Walter, J.C. Mechanism of RAD51-dependent DNA interstrand cross-link repair. *Science* **333**, 84–87 (2011).
12. Berezney, R., Dubey, D.D. & Huberman, J.A. Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. *Chromosoma* **108**, 471–484 (2000).
13. Kottemann, M.C. & Smogorzewska, A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature* **493**, 356–363 (2013).
14. Walden, H. & Deans, A.J. The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. *Annu. Rev. Biophys.* **43**, 257–278 (2014).
15. Williams, H.L., Gottesman, M.E. & Gautier, J. The differences between ICL repair during and outside of S phase. *Trends Biochem. Sci.* **38**, 386–393 (2013).
16. Bunting, S.F. & Nussenzweig, A. Dangerous liaisons: Fanconi anemia and toxic nonhomologous end joining in DNA crosslink repair. *Mol. Cell* **39**, 164–166 (2010).
17. Legerski, R.J. Repair of DNA interstrand cross-links during S phase of the mammalian cell cycle. *Environ. Mol. Mutagen.* **51**, 540–551 (2010).
18. Muniandy, P.A., Liu, J., Majumdar, A., Liu, S.-T. & Seidman, M.M. DNA interstrand crosslink repair in mammalian cells: step by step. *Crit. Rev. Biochem. Mol. Biol.* **45**, 23–49 (2010).
19. Nakanishi, K. *et al.* Homology-directed Fanconi anemia pathway cross-link repair is dependent on DNA replication. *Nat. Struct. Mol. Biol.* **18**, 500–503 (2011).
20. Willis, N.A. *et al.* BRCA1 controls homologous recombination at Tus/Ter-stalled mammalian replication forks. *Nature* **510**, 556–559 (2014).
21. Le Breton, C., Hennion, M., Arimondo, P.B. & Hyrien, O. Replication-fork stalling and processing at a single psoralen interstrand crosslink in *Xenopus* egg extracts. *PLoS ONE* **6**, e18554 (2011).
22. Huang, J. *et al.* The DNA translocase FANCM/MHF promotes replication traverse of DNA interstrand crosslinks. *Mol. Cell* **52**, 434–446 (2013).
23. Duxin, J.P., Dewar, J.M., Yardimci, H. & Walter, J.C. Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell* **159**, 346–357 (2014).
24. Crossan, G.P. & Patel, K.J. The Fanconi anaemia pathway orchestrates incisions at sites of crosslinked DNA. *J. Pathol.* **226**, 326–337 (2012).
25. Leonard, A.C. & Méchali, M. DNA replication origins. *Cold Spring Harb. Perspect. Biol.* **5**, a010116 (2013).
26. Duderstadt, K.E., Reyes-Lamothe, R., van Oijen, A.M. & Sherratt, D.J. Replication-fork dynamics. *Cold Spring Harb. Perspect. Biol.* **6**, a010157 (2014).
27. Blow, J.J., Ge, X.Q. & Jackson, D.A. How dormant origins promote complete genome replication. *Trends Biochem. Sci.* **36**, 405–414 (2011).
28. Ishibashi, T. & Lippard, S.J. Telomere loss in cells treated with cisplatin. *Proc. Natl. Acad. Sci. USA* **95**, 4219–4223 (1998).
29. Sobeck, A., Stone, S., Landais, I., de Graaf, B. & Hoatlin, M.E. The Fanconi anemia protein FANCM is controlled by FANCD2 and the ATR/ATM pathways. *J. Biol. Chem.* **284**, 25560–25568 (2009).
30. Branzei, D. & Foiani, M. Maintaining genome stability at the replication fork. *Nat. Rev. Mol. Cell Biol.* **11**, 208–219 (2010).
31. León-Ortiz, A.M., Svendsen, J. & Boulton, S.J. Metabolism of DNA secondary structures at the eukaryotic replication fork. *DNA Repair (Amst.)* **19**, 152–162 (2014).

ONLINE METHODS

LacI protein purification. The LacI-biotin protein was purified according to a protocol provided by the laboratory of K. Marians (Memorial Sloan Kettering Cancer Center). Briefly, NEB T7 express cells were cotransformed with pBirAcm, which contains an IPTG-inducible *birA* gene for overexpression of the biotin ligase (Avidity) and pET11a[lacI::avi] (a gift from the laboratory of K. Marians), containing the *lacI* gene encoding an AviTag at the C terminus. Transformed cells were grown on LB plates containing ampicillin and chloramphenicol, and single colonies were picked and amplified in LB containing the same antibiotics. Cells were lysed in lysis buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 1 mM DTT, 5 mM EDTA, 10% sucrose, 0.2 mg/ml lysozyme, 0.1% Brij 58 and cOmplete protease inhibitor (Roche)), and chromatin-bound LacI was extracted from the pellets by sonication. Nucleic acids were removed from the extracted fraction by addition of polymin P (final concentration 0.05%), and LacI-biotin was precipitated by addition of ammonium sulfate to a final concentration of 37%. The LacI-biotin pellet was then resuspended in 50 mM Tris, pH 7.5, 100 mM NaCl, 1 mM DTT, 1 mM EDTA and cOmplete protease inhibitor (Roche), applied to a soft-link avidin column and eluted with biotin-containing buffer (50 mM Tris, pH 7.5, 1 mM EDTA, 100 mM NaCl, 1 mM DTT and 5 mM biotin). LacI-biotin was dialyzed against 50 mM Tris, pH 7.5, 1 mM EDTA, 150 mM NaCl, 1 mM DTT, 38% glycerol overnight, frozen in liquid nitrogen and stored at −80 °C. A more detailed LacI-biotin purification protocol is available upon request.

Preparation of pICL-*lacO*. To make the backbone of pICL-*lacO*, we first engineered an EcoRI site 295 nt downstream of the second BbsI site in the parental plasmid of pICL[32]. We then cloned the *lacO* array (48 *lacO* repeats) between the EcoRI and SacI sites. The parental plasmid was then amplified and digested with BbsI. To make the cisplatin pICL-*lacO*, a 20-nt cisplatin-ICL duplex was prepared and ligated into the tandem BbsI sites of the backbone plasmid[32]. To make the trioxsalen ('psoralen' in the main text) pICL-*lacO*, complementary primers containing only one thymidine were annealed in annealing buffer (100 mM potassium acetate, 30 mM HEPES-KOH, pH 7.4, and 2 mM magnesium acetate) at a concentration of 50 μM each. DNA-trioxsalen cross-linking was carried out with 2.6 μM annealed DNA in cross-linking buffer (10 mM Tris, pH 7.5, 1 mM EDTA, and 50 mM NaCl) and 87.6 μM trioxsalen. The reaction was exposed to 365-nm UVA light for six periods of 15 min each, at a power of 4 mW/cm[2]. After every cycle, fresh trioxsalen was added to 87.6 μM. Cross-linked DNA was purified from a 20% polyacrylamide 8 M urea gel. The purified DNA was end-labeled with [γ-^{32}P]ATP and run on a gel, which revealed that >99.9% of the DNA contained a DNA interstrand cross-link. The purified cross-linked duplex was then ligated into the tandem BbsI sites of its corresponding backbone plasmid. Sequences were as follows: cisplatin cross-linked duplex (with the cross-link between the two Gs in bold), 5′-CCCTCCTTCC**G**CTCCTTCTTTC-3′ and 5′-GCACGAAAGAAGA**G**CGGAAG-3′; psoralen cross-linked duplex (with the cross-link between the two Ts in bold): 5′-CCCCGGGGC**T**AGCC-3′ and 5′-GCACGGC**T**AGCCCC-3′.

The generation of the *lacO* array (48 *lacO* repeats) will be presented in detail elsewhere (J.M.D. and J.C.W., unpublished data). The sequences of the plasmids and primers used for mutagenesis are available upon request.

***Xenopus* egg extracts and DNA replication.** *Xenopus* egg extracts were prepared as previously described[33]. For DNA replication, plasmids were first incubated in a high-speed supernatant (HSS) of egg cytoplasm (final concentration of 7.5 ng DNA/μL extract) for 20 min at room temperature to license the DNA; this was followed by the addition of two volumes of nucleoplasmic egg extract (NPE) to initiate replication. For replication with LacI, plasmid (75 ng/μL) was incubated with an equal volume of 40 μM LacI for 30 min before HSS addition. IPTG was added at a final concentration of 10 mM in egg extracts. In **Supplementary Figure 3c**, pICL-*lacO* was incubated in HSS without prebinding of LacI, and LacI was added at the indicated times relative to NPE addition. In all figures, the 0-min time point refers to the time of NPE addition. For DNA labeling, reactions were supplemented with [α-^{32}P]dATP, which is incorporated into nascent strands during replication. Where indicated, pQuant, an undamaged plasmid lacking the *lacO* array, was included (0.375 ng/μL final concentration in HSS) as an internal replication standard. For **Figure 3a**, replication was stopped by addition of 0.5 μl of each reaction to 10 μl of replication stop solution A (5% SDS, 80 mM Tris, pH 8.0, 0.13% phosphoric acid and 10% Ficoll) supplemented with 1 μl Proteinase K (20 mg/ml) (Roche). Samples were incubated for 1 h at 37 °C before separation by 0.8% native agarose gel electrophoresis. DNA samples were then detected with

a phosphorimager[33]. For all other applications except ChIP, replication reactions were stopped in ten volumes of replication stop solution B (50 mM Tris, pH 7.5, 0.5% SDS and 25 mM EDTA), and replication intermediates were purified as previously described[6]. Experiments in **Figures 1d** and **3** were performed three times; experiments in **Figures 1f** and **4c,d** were performed twice; experiment in **Figure 2b** was performed four times, and representative results are shown.

ChIP and quantitative real-time PCR. ChIP was performed essentially as previously described[11]. Briefly, 3-μL reaction samples were cross-linked in 47 μL of 1% formaldehyde in ELB (10 mM HEPES-KOH, pH 7.7, 2.5 mM MgCl$_2$, 50 mM KCl and 250 mM sucrose) for 10 min at room temperature. Cross-linking was stopped by the addition of 5 μL 1.25 M glycine and subsequent passage through a Micro Bio-Spin 6 chromatography column (Bio-Rad) to remove excess formaldehyde. The flow-through was diluted to 500 μL with sonication buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 2 mM EDTA, 0.5% NP-40, 5 μg/mL aprotinin plus leupeptin and 2 mM PMSF) and subjected to sonication, yielding DNA fragments ~300–500 bp in size.

After immunoprecipitation (IP), formaldehyde cross-links were reversed, and DNA was purified for analysis by quantitative real-time PCR. The recovery rate was determined by the amount of IP samples relative to the input samples. For quantitative PCR, in which three technical replicates were performed for each sample, replicates that deviated from the average value by greater than 0.3 s.d. were discarded. Polyclonal antibodies to FANCD2 (ref. 6; rabbit 20019), CDC45 (ref. 34; rabbit 534), MCM7 (ref. 35; rabbit 456), XPF (ref. 10; rabbit 20682 and rabbit 20683), and SLX4 (ref. 10; rabbit 24153 and rabbit 24256) were previously described and validated. For MCM7, SLX4 and FANCD2 IP, antibodies were purified from serum with Protein A–Sepharose beads (GE Healthcare), and 5 μg of IgG was used for IP per sample. For CDC45 and XPF, 1 μL of serum was used directly for immunoprecipitation.

PCR primer pairs used in ChIP are as follows: ICL primer pair, 5′-AGCC AGATTTTTCCTCCTCTC-3′ and 5′-CATGCATTGGTTCTGCACTT-3′; control primer pair on pICL-*lacO*, 5′-AACGCCAATAGGGACTTTCC-3′ and 5′-GGGCGTACTTGGCATATGAT-3′; control primer pair on pQuant, 5′-TA CAAATGTACGGCCAGCAA-3′ and 5′-GAGTATGAGGGAAGCGGTGA-3′.

Nascent-strand analysis. Nascent-strand analysis was performed as previously described[6]. Briefly, pICL-*lacO* was replicated in the presence of [α-^{32}P]dATP, and purified repair intermediates were digested with AflIII and EcoRI; this was followed by addition of 0.5 volumes gel loading buffer II (denaturing PAGE) (Life Technologies). For the mixing experiment of pICL and p*lacO* in **Supplementary Figure 3d**, the repair intermediates were digested with only AflIII, which cuts on both sides of the ICL on pICL. Radiolabeled nascent strands were then separated on a 7% denaturing polyacrylamide gel, transferred to filter paper, dried and visualized with a phosphorimager. Sequencing gel markers were generated with primer S (5′-CATGTTTTACTAGCCAGATTTTTCCTCCTCTCCTG-3′) with the Cycle Sequencing kit (USB Corporation).

Incision assay. The incision assay was performed as described before[36]. Briefly, pICL-*lacO* was replicated and digested with BlpI, which cuts the plasmid once. In parallel, unreplicated pICL-*lacO* was digested with BlpI to serve as size markers for the X-shaped structure and linear DNA, which came from a small fraction of un-cross-linked plasmids present in our pICL-*lacO* preparations. After digestion, the DNA was separated on a 1% agarose gel under denaturing conditions (50 mM NaOH and 1 mM EDTA) for 18 h at 0.85 V/cm. Subsequently, Southern blotting was performed by capillary transfer in transfer buffer (1.5 M NaCl and 0.4 M NaOH) onto a nylon membrane (Hybond-N+, Amersham). After transfer, the membrane was washed in 4× SSC for 5 min and was UV irradiated to cross-link the DNA to the membrane. Prehybridization was performed with 25 ml of hybridization buffer (4× SSC, 2% SDS, 1× blocking reagent (Roche), 0.1 mg/ml salmon sperm DNA (Life Technologies)) for 30 min at 45 °C. Hybridization was carried out overnight with 25 μl of probe prepared with a Roche random labeling kit (Roche). After overnight hybridization, the membrane was washed four times with 0.5× SSC and 0.25% SDS for 15 min at 45 °C. The dried membrane was exposed to a phosphorimager screen.

Strand-specific Southern blot. Strand-specific Southern blot was performed as previously described[6]. Briefly, AflIII- and AseI-digested samples were separated

on a 7% polyacrylamide gel and transferred to a nylon membrane (Hybond-N+, Amersham). After transfer, the membrane was rinsed in 4× SSC for 5 min and was UV irradiated to cross-link the DNA to the membrane. The membrane was then prehybridized with 25 ml hybridization buffer (ULTRAhyb, Ambion) for at least 3 h at 42 °C. Strand-specific probes generated by a PCR-based primer-extension reaction[6] were added to the hybridization buffer and incubated with the membrane at 42 °C overnight. After overnight hybridization, the membrane was washed two times with 2× SSC and 0.1% SDS for 5 min at 42 °C. The dried membrane was exposed to a phosphorimager screen.

Uncropped images of gels and autoradiographs used in this study can be found in **Supplementary Data Set 1**.

32. Enoiu, M., Ho, T.V., Long, D.T., Walter, J.C. & Schärer, O.D. Construction of plasmids containing site-specific DNA interstrand cross-links for biochemical and cell biological studies. *Methods Mol. Biol.* **920**, 203–219 (2012).
33. Lebofsky, R., Takahashi, T. & Walter, J.C. DNA replication in nucleus-free *Xenopus* egg extracts. *Methods Mol. Biol.* **521**, 229–252 (2009).
34. Walter, J. & Newport, J. Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase α. *Mol. Cell* **5**, 617–627 (2000).
35. Fang, F. & Newport, J.W. Distinct roles of cdk2 and cdc2 in RP-A phosphorylation during the cell cycle. *J. Cell Sci.* **106**, 983–994 (1993).
36. Knipscheer, P., Räschle, M., Schärer, O.D. & Walter, J.C. Replication-coupled DNA interstrand cross-link repair in *Xenopus* egg extracts. *Methods Mol. Biol.* **920**, 221–243 (2012).
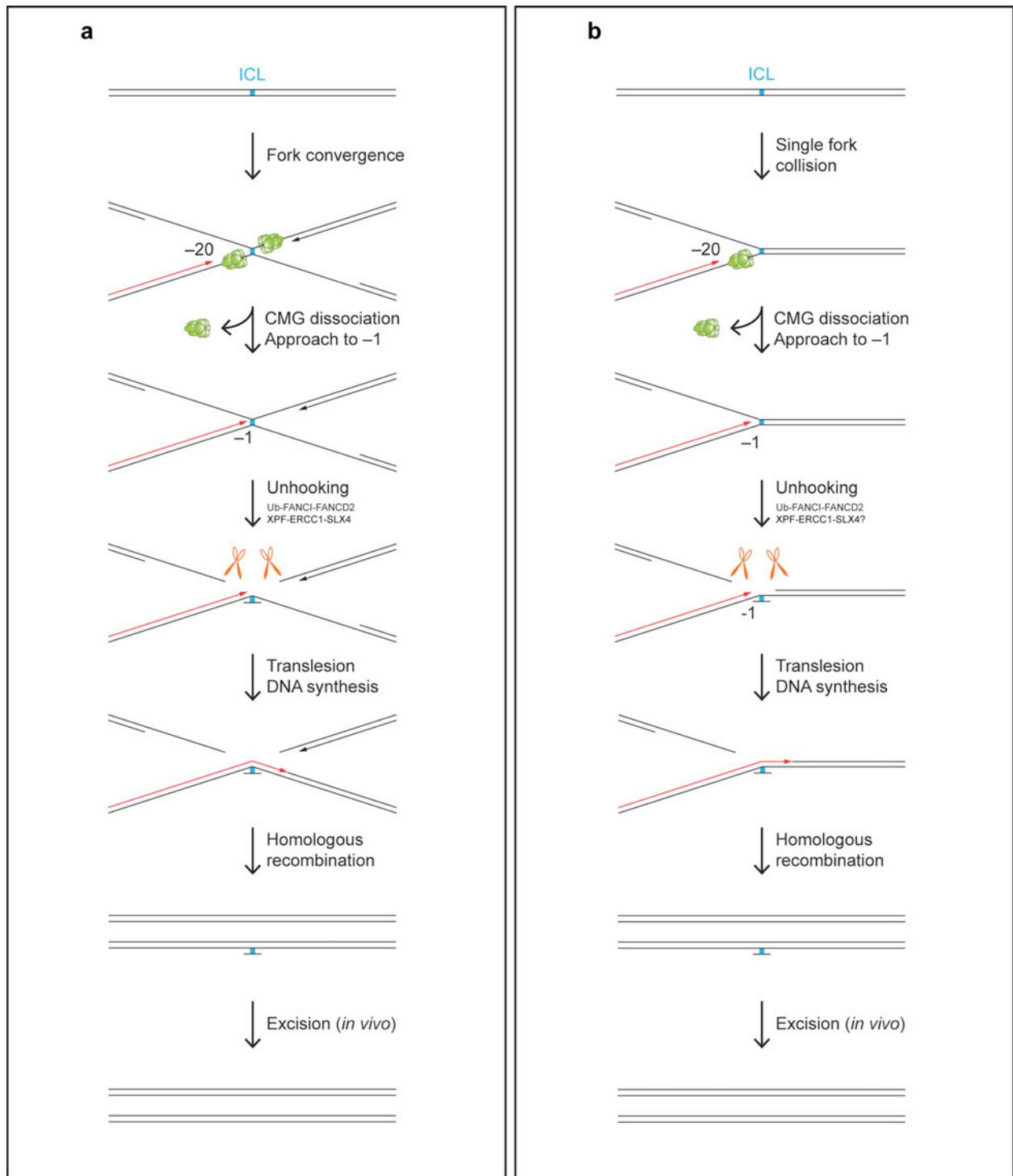
Supplementary Figure 1

Model for ICL repair with two forks and one fork.

Model for ICL repair with two forks (a) and one fork (b). The details of the single fork model are inferred from the mechanism previously established for the fork convergence model[6-11].
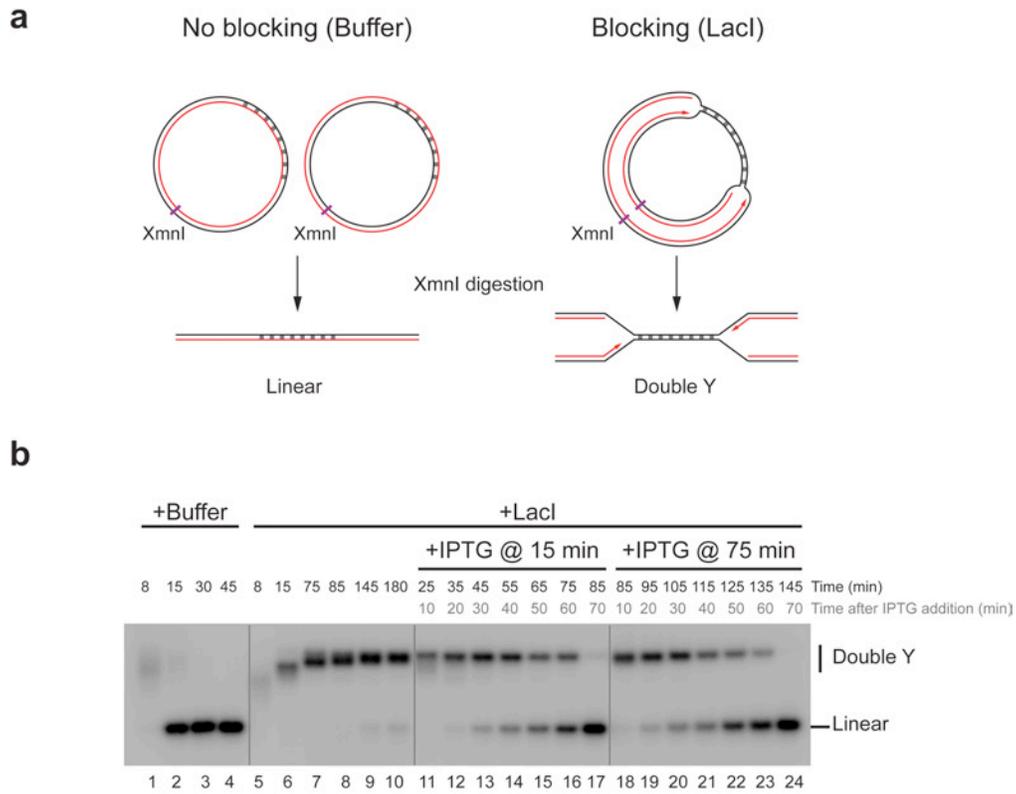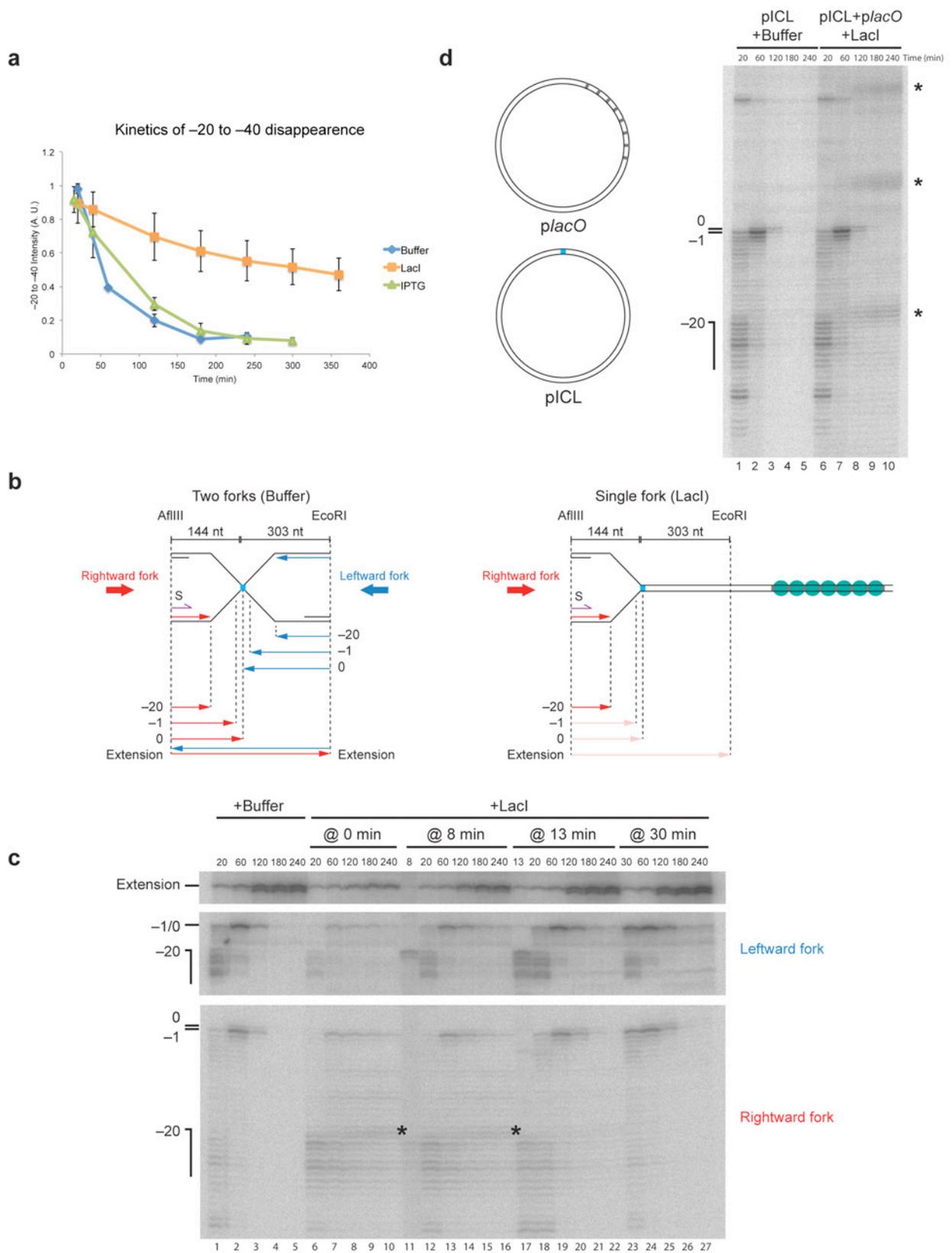
Supplementary Figure 2

The LacI-*lacO* array efficiently blocks fork progression.

(a) Schematic of p*lacO* replication intermediates expected in the presence and absence of LacI after digestion with XmnI.

(b) p*lacO* was replicated with or without LacI, and IPTG was added at the indicated times. Replication intermediates were digested with XmnI, separated on an agarose gel, and detected by autoradiography.

a Kinetics of −20 to −40 disappearence

b Two forks (Buffer) / Single fork (LacI)

c +Buffer / +LacI

d placO / pICL

Supplementary Figure 3

Leading strands persist at –20 to –40 when there is only a single fork.

(a) The intensity of leading strands located between the –20 to –40 positions in Fig. 1d and two repetitions of this experiment was quantified and graphed. Error bars represent standard deviations.

(b) Schematic illustration of nascent leading strands generated in the experiment described in (c).

(c) The LacI-*lacO* array does not inhibit ICL repair once forks have converged on the ICL. pICL-*lacO* was replicated in egg extracts, and LacI was added at the indicated times relative to NPE addition. Nascent leading strands were digested with AflIII and EcoRI and analyzed on a sequencing gel, as described in Fig. 1d. Note that LacI addition at 0 and 8 minutes after replication initiation, when most forks had not yet converged on the ICL (lane 11), inhibited approach (lanes 6–10 and 12–16). In contrast, LacI addition at 13 and 30 minutes, when most forks had converged (lanes 17 and 23), was not inhibitory for approach (lanes 18–22 and 24–27). Asterisk (*), background bands described in (d).

(d) The LacI-*lacO* array does not inhibit the approach of pICL repair *in trans.* pICL (which lacks the *lacO* array) and p*lacO* (no ICL) were mixed and replicated in the presence of buffer or LacI, as indicated. Samples were digested with AflIII, and leading strands of pICL were monitored on a sequencing gel. A series of species differing in size by multiples of 30 nt was observed whenever a *lacO*-containing plasmid was replicated in the presence of LacI (* bands in the right panel, see also Fig. 1d, 1f; Supplementary Figure. 3c, 4b). The formation of the ladder was independent of restriction enzyme digestion (data not shown), indicating that non-ligated nascent strands generated within the *lacO* array give rise to the repeating pattern.
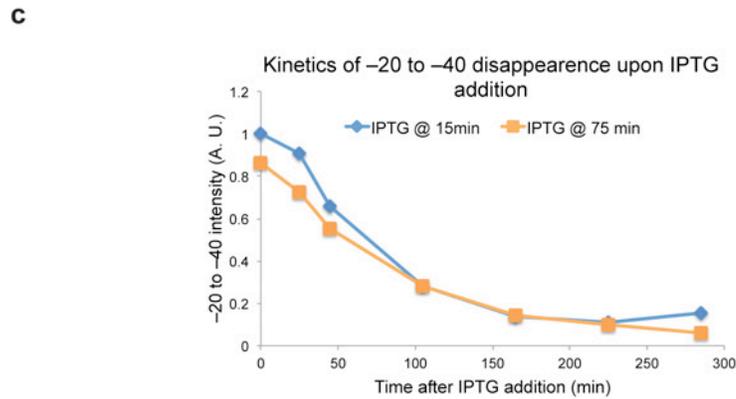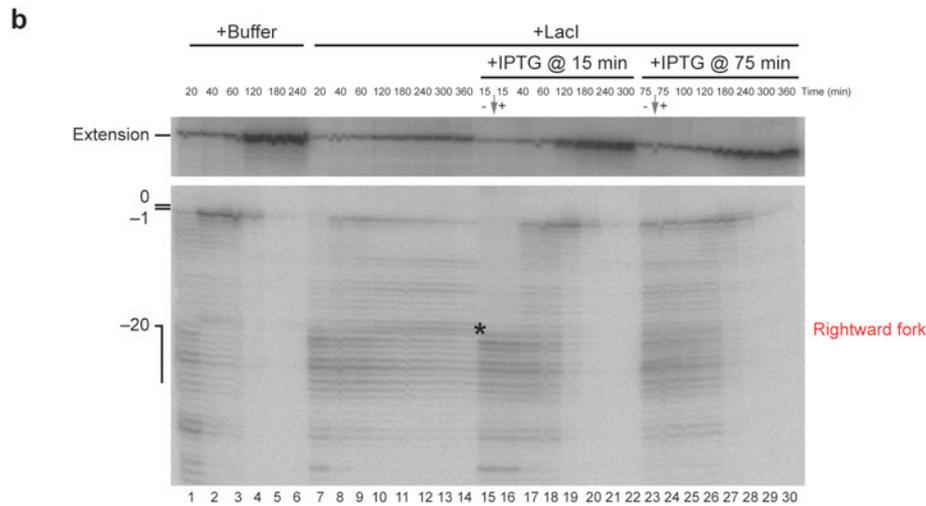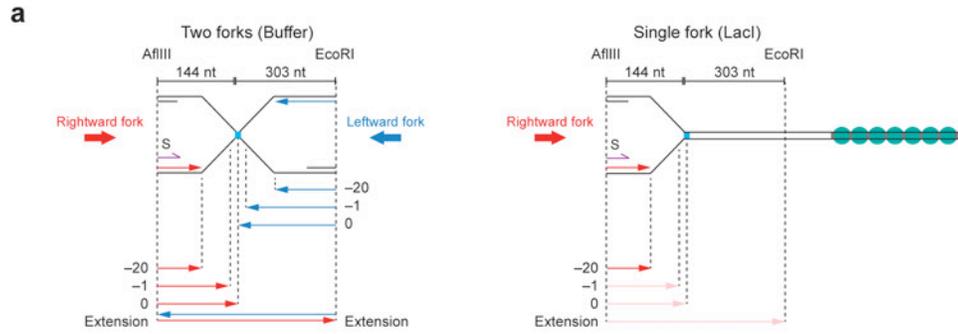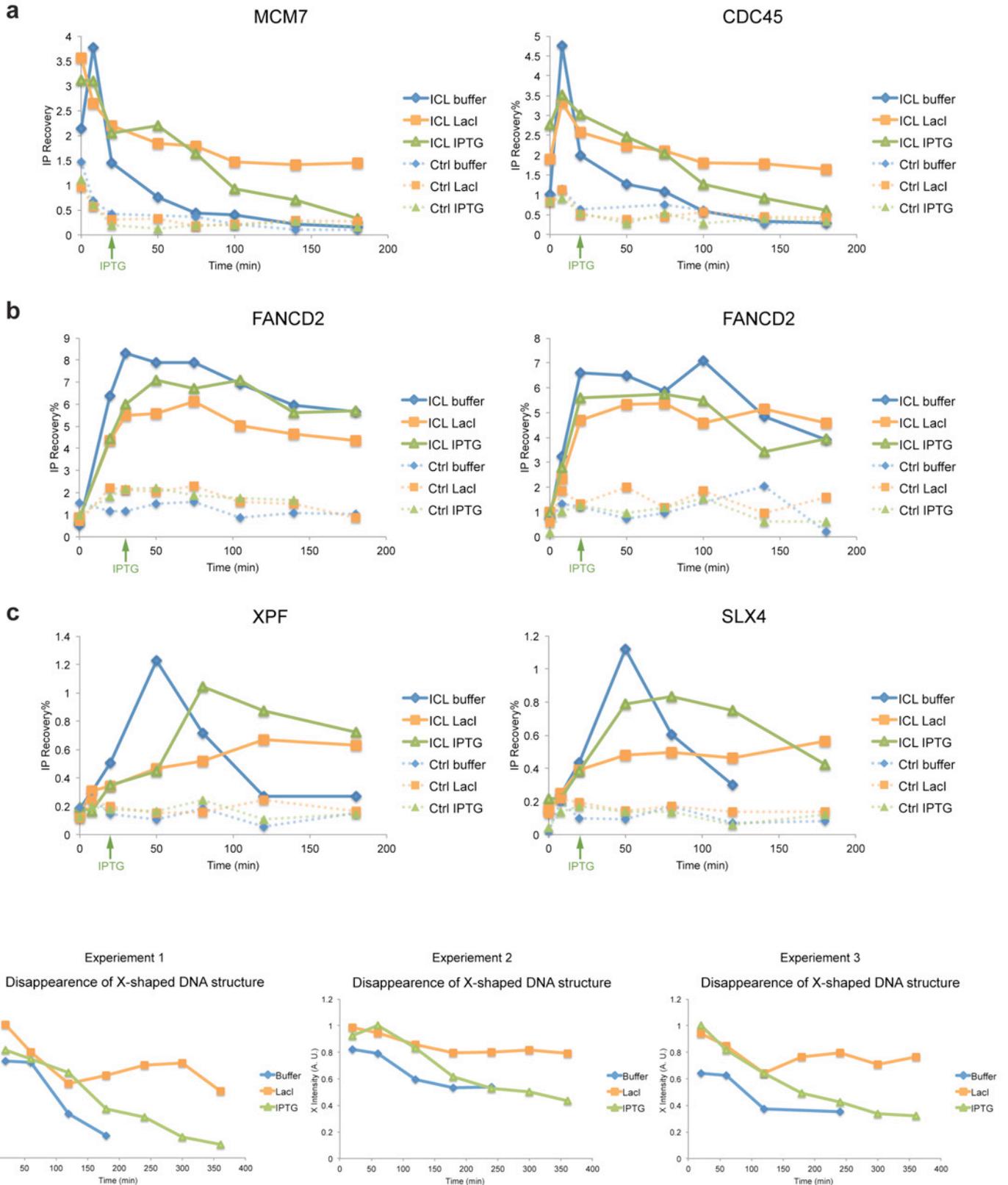
Supplementary Figure 4

A single fork remains competent for ICL repair after prolonged stalling.

(a) Schematic illustration of nascent leading strands generated in the presence and absence of LacI.

(b) pICL-*lacO* was replicated with or without LacI, and IPTG was added at the indicated times. Nascent leading strands were digested with AflIII and EcoRI and analyzed on a sequencing gel, as described in Fig. 1d. Asterisk (*), background bands described in Supplementary Fig. 3d.

(c) The intensity of the –20 to –40 products after IPTG addition in (b) was quantified and graphed.

Supplementary Figure 5

ChIP results and quantification of incision assays.

**(a-c) ChIP results for MCM7, CDC45, FANCD2, XPF, and SLX4**

(a) Experimental replicate of the MCM7 and CDC45 ChIP described in Fig. 2b.

(b) Two independent examples of FANCD2 ChIP. pICL-*lacO* and pQuant were replicated with or without LacI, and IPTG was added immediately before the 20 minute time point, as indicated (green arrow). At different times, samples were withdrawn for FANCD2 ChIP using primer pairs for the ICL locus (Fig. 2a) or pQuant (Ctrl).

(c) Experimental replicate of the XPF and SLX4 ChIP described in Fig. 3e.

**(d) Quantification of Incision Assay**

We quantified and graphed the efficiency of incision in the experiment shown in Fig. 3d (left) and in two repetitions of this experiment (middle and right). Incisions convert the X-shaped parental structure into linear products (Fig. 3c, left). However, linear products are created not only via incision, but also from nascent strands generated during lesion bypass, from homologous recombination, and from replication of undamaged plasmid. In contrast, the disappearance of X-shaped molecules is caused exclusively by incision, making the reduction in the intensity of the X-shaped band a better readout of incision. To compare incisions in the presence of one and two forks, we therefore determined the reduction of the X-shaped species in the presence of buffer (two forks), in the presence of LacI (one fork), and in the presence of LacI and IPTG (two forks). We found that in the three experiments shown above, LacI inhibited the reduction in X-shaped species by an average of ~70%, consistent with the fact that on 26% of molecules, a leftward replication fork reaches the ICL even in the presence of LacI (Fig. 1d).
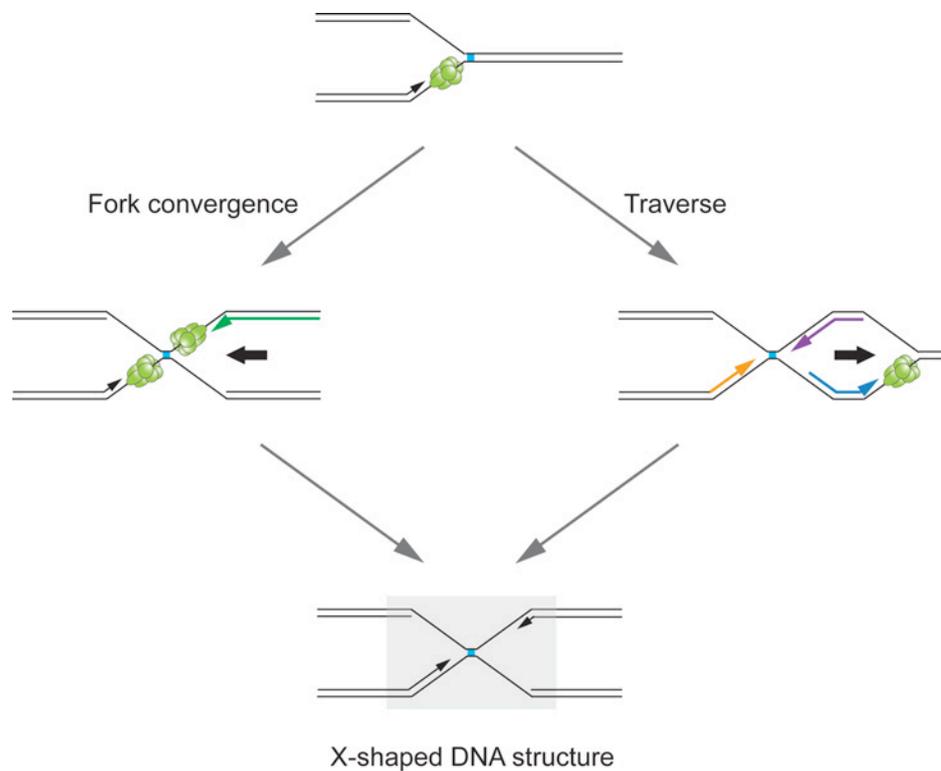
Supplementary Figure 6

Fork convergence versus traverse.

When a single DNA replication fork encounters an ICL, it can pause until the arrival of a second fork (left arrow, "fork convergence") or bypass the lesion (right arrow, "traverse")[22]. In both cases, an X-shaped DNA structure surrounding the ICL is generated (grey box). If traverse happened in our system, the leading strands of the rightward, single fork should approach to –1 (orange strand) since CMG either dissociates or bypasses the ICL, yet they stall at the –20 position. The 5′ ends of the leading strands of the traversed fork might be heterogeneous, depending on where DNA synthesis restarts on the other side of the ICL (blue strand). In addition, lagging strands of the traversed fork should approach directly to the lesion without –20 stalling since there is no CMG travelling ahead of the polymerase (purple strand). Although we observed a low abundance of nascent strands between the ICL and the *lacO* array, they initially stalled at the –20 position, after which they underwent approach (Fig. 1d, 1f, leftward fork panel). Therefore, we conclude that these signals came from the arrival of the leftward forks (green strand; Supplementary Figure. 2) rather than traverse (purple strand).
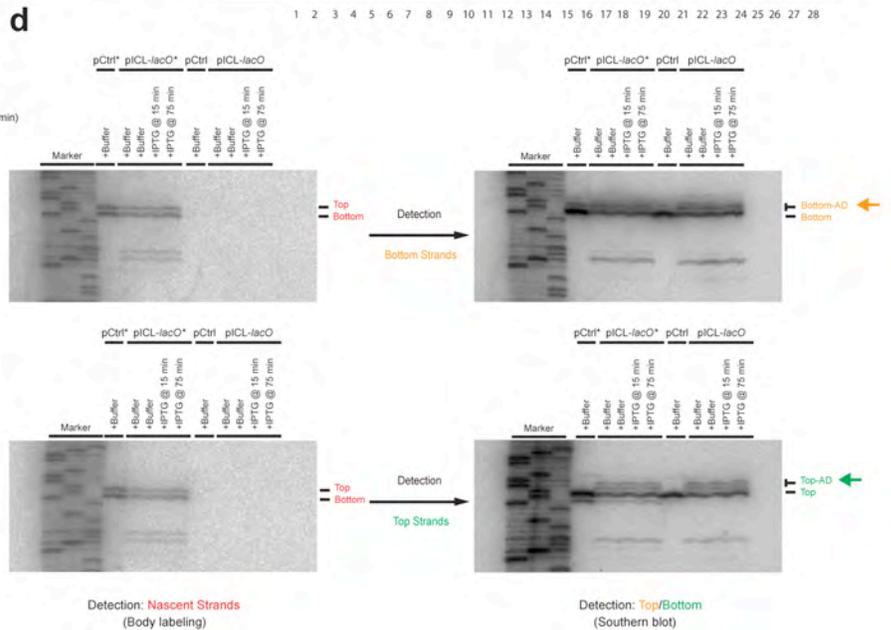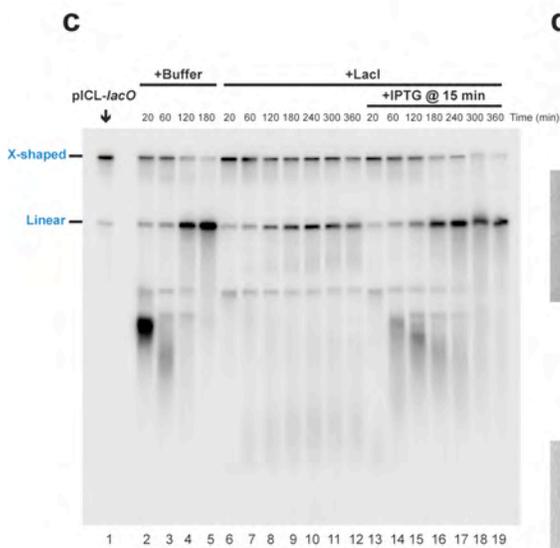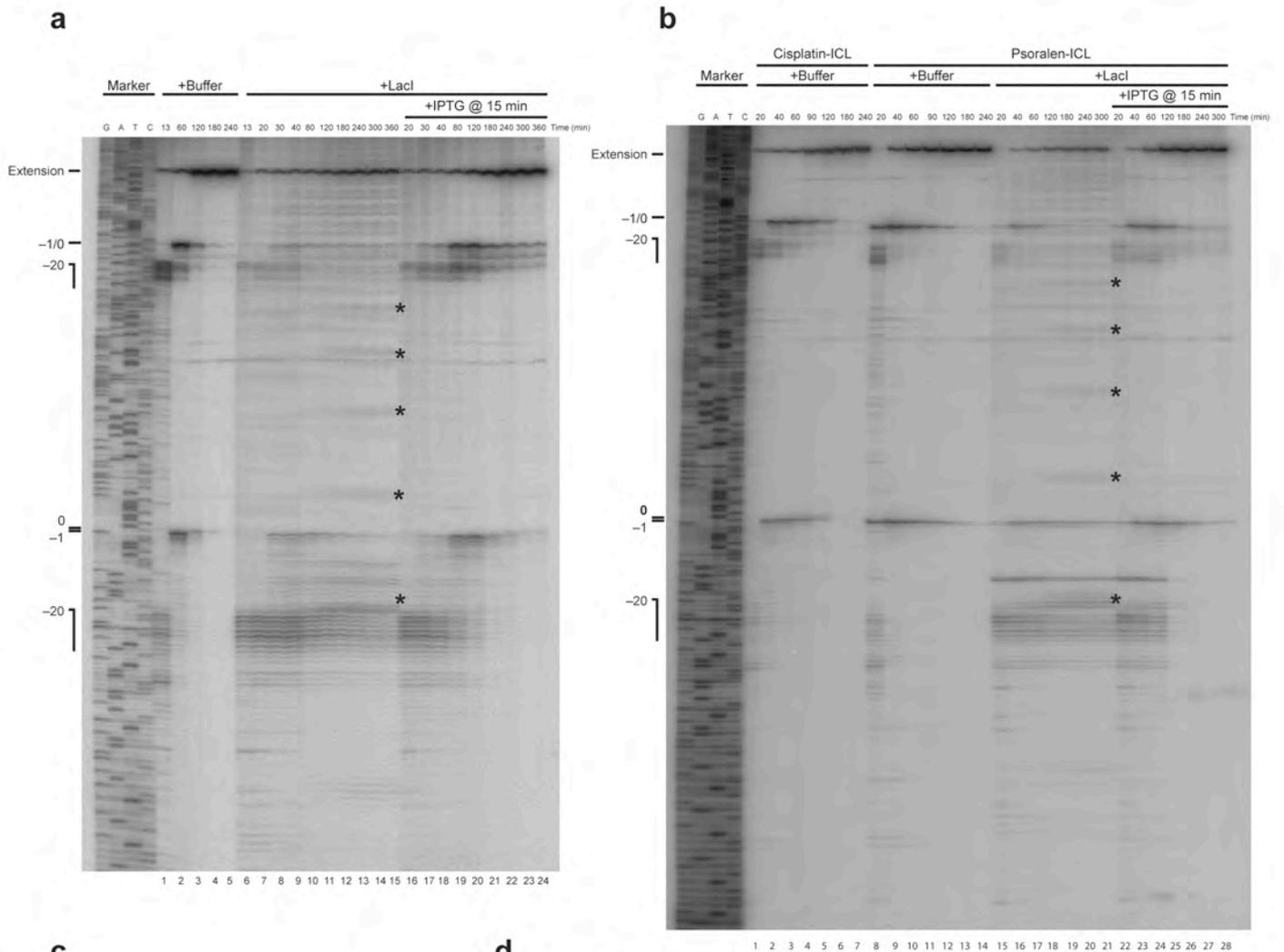
Supplementary Data Set 1

Uncropped images of main figures

(a) Uncropped image of Fig. 1d. Asterisk (*), background bands described in Supplementary Fig. 3d.

(b) Uncropped image of Fig. 1f. Asterisk (*), background bands described in Supplementary Fig. 3d.

(c) Uncropped image of Fig. 3d. The unlabeled smaller fragments running below the linear species correspond to incision products or the non-ligated nascent strands.

(d) Uncropped images of Figs. 4c and 4d.

pCtrl or pICL-*lacO* was replicated in the presence or absence of [α-^{32}P]dATP (*), and buffer/LacI/IPTG was added as indicated. Before Southern blotting probe was added (left panels), the nascent strands were visualized by autoradiography and served as size markers for top/bottom strands illustrated in Fig.4b. Note that the signals were detected only when [α-^{32}P]dATP was present during DNA replication. Besides the expected top and bottom bands, a low level of smaller band was also observed, corresponding to deletion mutants generated during repair that will be discussed elsewhere. After strand-specific Southern blotting (right panels), we were able to detect almost exclusively either the top or bottom strands. Note that when [α-^{32}P]dATP was present, the total signal reflected the combination of radioactive nascent strands and Southern blotting probe. However, when [α-^{32}P]dATP was omitted during replication, all the radioactive signal came from Southern blotting.