Molecular Mechanisms of DNA Interstrand Cross-Link Repair by the Fanconi Anemia Pathway

Thesis by

Maria Altshuller

In Partial Fulfillment of the Requirements for the

Degree of

Doctor of Philosophy in Biochemistry and Molecular Biophysics



CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2026

(Defended May 21st, 2026)

© 2026

Maria Altshuller
ORCID: 0009-0006-8174-7713

All rights reserved except where otherwise noted

ACKNOWLEDGEMENTS

I am deeply grateful to my advisor, Prof. Dan Semlow, whose mentorship has profoundly influenced my growth as a researcher. Not only have his insight and creativity shaped the way I approach scientific problems, but he has also been unfailingly supportive and kind throughout my years in his lab, which I appreciate more than I can express. I could not have asked for a better environment for my graduate studies.

I would next like to thank the other members of my thesis committee, Prof. Shu-ou Shan, Prof. Bill Dunphy, Prof. Mitch Guttman, and Prof. Rebecca Voorhees, for their guidance and thoughtful feedback throughout my graduate years. Their perspectives have consistently challenged and refined my thinking, and this thesis is all the better for their engagement with it. I would also like to extend special thanks to Prof. Judy Campbell, who is an invaluable wellspring of knowledge in our field, and whose generous encouragement over the years has meant a great deal to me.

I also want to thank the current and former members of the Semlow Lab: Shay He, Jasmine Tzeng, Terry Kim, Shuangshuang Xie, Richa Nigam, Matt Midy, Gabriel Ashton-Rickardt, Omar Salah, Jade Konsler, and Charlotte Chai. The laughter and camaraderie you have all brought to our lab have been a constant source of joy throughout my PhD. I also want to extend thanks to Ann Mao and Rebecca Fox for their help with countless administrative tasks, as well as for their warmth and support through all my years at Caltech.

I am especially grateful to my longtime Semlow Lab baymate and dearest friend, Victoria MacKrell. Her brilliant scientific vision has contributed to my work in countless ways, and

her wisdom and steadfast support have been a source of strength I could not have done without. I have gained many things from my graduate years, but her friendship is the most cherished of them. I also want to thank Elliot MacKrell, a wonderful friend whose warmth and good humor have brightened many a difficult day, and their son Simon MacKrell, who is the textbook definition of a bundle of joy.

Graduate school would have been a much lonelier experience without my fellow PhD students. I am particularly grateful to Kate Radford, Noah Epstein, Richard Horak, Will Rosencrans, and Dan Pollak, whose companionship, humor, and warmth have enriched my life outside the lab and kept me sane throughout my PhD. I also want to thank the members of OcTech, the choral octet I have been lucky to be a part of during my time at Caltech: Kate, Philip, Georgia, Noah, James, Alma, Hannah, Magdi, and Michelle. Singing with you every week has brought me enormous happiness.

Lastly, I want to thank my parents, Irina and Vladlen, whose love and unwavering support are the foundation on which everything else is built.

ABSTRACT

DNA interstrand cross-links (ICLs), which consist of a covalent link between complementary DNA strands, are among the most cytotoxic lesions that cells encounter, as the ICL blocks strand separation during replication and transcription. The Fanconi anemia (FA) pathway is the primary mechanism through which proliferating cells repair ICLs encountered during DNA replication, and mutations in FA pathway genes cause the inherited bone marrow failure and cancer predisposition syndrome Fanconi anemia. Despite considerable progress in elucidating the molecular mechanism of FA pathway-mediated ICL repair, fundamental questions remain about how specific ICL substrates are processed and how the replication fork dynamics that accompany repair are regulated. This thesis addresses these questions using cell-free extracts prepared from *Xenopus laevis* eggs, which enable biochemical characterization of ICL repair. We first investigated the repair of ICLs induced by colibactin, a gut microbiome-derived genotoxin implicated in the pathogenesis of colorectal cancer. We demonstrate that the FA pathway is responsible for replication-coupled repair of colibactin-induced ICLs and identify the translesion synthesis polymerases that bypass the unhooked ICL remnant, providing mechanistic insight into how cells process this clinically relevant lesion. We then uncovered a phosphoregulatory axis governing replication fork dynamics during ICL repair, showing that the ataxia-telangiectasia mutated (ATM) kinase promotes nucleolytic resection of reversed fork intermediates through the nucleases EXO1 and DNA2-WRN, and that this resection is constrained by opposing PP2A phosphatase activity. This regulatory mechanism ensures productive fork restoration. Finally, we present preliminary evidence that X-shaped replication intermediates are sufficient to activate FA pathway signaling and ICL unhooking in the absence of ongoing replication, suggesting that these structures may serve as the trigger for FA pathway activation independently of the replisome. Taken together, this work advances our mechanistic understanding of how cells tolerate ICL stress and illuminates regulatory mechanisms that ensure accurate and efficient ICL repair.

PUBLISHED CONTENT AND CONTRIBUTIONS

1 Altshuller, M.; He, X.; MacKrell, E. J.; Wernke, K. M.; Airan, Y.; Wong, J. W. H.; Sellés-Baiget, S.; Wang, T.-Y.; Chou, T.-F.; Duxin, J. P.; Balskus, E. P.; Herzon, S. B.; Semlow, D. R. The Fanconi anemia pathway repairs colibactin-induced DNA interstrand cross-links. *Nat. Commun.* **16**, 10563 (2025). https://doi.org/10.1038/s41467-025-65606-1

M.A. participated in the conception of the project, collection and analysis of data, and preparation of the manuscript. See Chapter 2.7 for details.

2 Altshuller, M.; MacKrell, V. A.; Tzeng, J.; Nigam, R.; Konsler, J. K.; Wang, T.-Y.; Grogan, E. J.; Ozua, O. E.; Campbell, J. L.; Dewar, J. M.; Chou, T.-F.; Semlow, D. R. ATM promotes nucleolytic processing of reversed replication forks during DNA interstrand cross-link repair. *bioRxiv* (2025). https://doi.org/10.1101/2025.10.08.680157

M.A. participated in the conception of the project, collection and analysis of data, and preparation of the manuscript. See Chapter 3.7 for details.

TABLE OF CONTENTS

Acknowledgements iii

Abstract v

Published Content and Contributions vi

Table of Contents vii

List of Illustrations x

List of Tables xii

Nomenclature xiii

Chapter 1: Introduction 1

1.1 DNA Interstrand Cross-Links: Causes and Consequences 1

1.2 Mechanisms of ICL Repair 2

1.3 The Fanconi Anemia Pathway 4

1.4 Colibactin: A Microbiome-Derived ICL-Inducing Genotoxin 6

1.5 The *Xenopus* Egg Extract System 7

1.6 Open Questions 10

1.7 Summary of This Thesis 11

1.8 References 13

Chapter 2: The Fanconi Anemia Pathway Repairs Colibactin-Induced DNA Interstrand Cross-Links 21

2.1 Abstract 22

2.2 Introduction 23

2.3 Results 28

2.4 Discussion 51

2.5 Methods 57

2.6 Acknowledgements 78

2.7 Author Contributions 79

2.8 Declaration of Interests 79

2.9 References 80

2.10 Supplementary Information 90

Chapter 3: ATM Promotes Nucleolytic Processing of Reversed Replication Forks During DNA Interstrand Cross-Link Repair 108

3.1 Abstract 109

3.2 Introduction 110

3.3 Results 114

3.4 Discussion 132

3.5 Methods 137

3.6 Acknowledgements 155

3.7 Author Contributions 155

3.8 Declaration of Interests 156

3.9 References 156

3.10 Supplementary Information 169

Chapter 4: Replication-Independent Activation of the Fanconi Anemia Pathway by DNA Replication Intermediates 180

4.1 Abstract 181

4.2 Introduction 182

4.3 Preliminary Data 183

4.4 Future Directions 187

4.5 Methods 193

4.6 References 196

LIST OF ILLUSTRATIONS

Number *Page*

2.1 Replication-coupled repair of colibactin-induced ICLs....24

2.2 FA pathway mobilization during pICL^{c742} replication 33

2.3 FA pathway-dependent unhooking of colibactin-induced ICLs....36

2.4 Translesion synthesis during colibactin-induced ICL repair 39

2.5 Translesion synthesis during bypass of a colibactin-induced monoadduct 43

2.6 Mutagenic repair of colibactin-induced ICLs by the FA pathway 45

2.7 REV1 and Pol κ promote mutagenic bypass of colibactin-induced DNA damage 48

2.8 Model for repair of colibactin-induced DNA damage....52

S2.1 Preparation and replication of plasmids containing colibactin-induced ICLs 91

S2.2 Mass spectrometry analyses of c742-ICL repair 95

S2.3 Colibactin- and cisplatin-induced ICLs are unhooked by the FA pathway 97

S2.4 Pol η, REV1, and Pol κ promote translesion synthesis during colibactin-induced ICL repair....100

S2.5 Pol κ and REV1 promote translesion synthesis past a colibactin-induced monoadduct....102

S2.6 Independent experimental replicates demonstrating mutagenic repair of colibactin-induced ICLs by the FA pathway 103

S2.7 Next-generation sequencing shows the mutagenicity of TLS past c742- and clb-induced adducts 105

3.1 Genome-wide CRISPR dropout screens implicate PTPA in DNA interstrand cross-link repair 111

3.2 Phosphatase inhibition induces replication fork collapse during ICL repair.....117

3.3 ATM promotes reversed fork processing during ICL repair 122

3.4 ATM activation during ICL repair requires CMG unloading.....127

3.5 EXO1 and DNA2-WRN mediate processing of reversed fork intermediates during ICL repair.....129

3.6 Model for regulation of reversed fork processing by ATM.....133

S3.1 Validation of RPE-1 cells used in CRISPR dropout screen 170

S3.2 Characterization of okadaic acid-induced fork collapse products 171

S3.3 Effects of CDK and DNA damage checkpoint inhibitors on ICL repair by the FA pathway 174

S3.4 ATM activation promotes processing of reversed fork intermediates.....176

S3.5 Nuclease-dependent processing of reversed fork intermediates.....178

4.1 Purified X-structures undergo ICL resolution and activate FA pathway signaling in non-replicating extract 184

4.2 Model for FA pathway-mediated ICL unhooking following fork convergence or ICL traverse 189

LIST OF TABLES

Number *Page*

S2.1 Oligonucleotide sequences used in this study.....90

S3.1 Oligonucleotide sequences used in this study.....169

NOMENCLATURE

AP site. Abasic (apurinic/apyrimidinic) site.

ATM. Ataxia-telangiectasia mutated kinase.

ATR. Ataxia telangiectasia and Rad3-related kinase.

CDC7. Cell division cycle 7 kinase.

CHK1. Checkpoint kinase 1.

CHK2. Checkpoint kinase 2.

clb-ICL. Colibactin-induced interstrand cross-link.

CMG. CDC45-MCM2-7-GINS replicative helicase.

CRC. Colorectal cancer.

CRISPR. Clustered regularly interspaced short palindromic repeats.

DNA2. DNA replication helicase/nuclease 2.

DSB. DNA double-strand break.

EXO1. Exonuclease 1.

FA. Fanconi anemia.

FANC. Fanconi anemia complementation group.

FDR. False discovery rate.

HR. Homologous recombination.

HSS. High-speed supernatant.

ICL. Interstrand cross-link.

LFQ. Label-free quantification.

MMR. Mismatch repair.

MRN. MRE11-RAD50-NBS1 complex.

NER. Nucleotide excision repair.

NGS. Next-generation sequencing.

NPE. Nucleoplasmic extract.

OC. Open circular.

PIKK. Phosphatidylinositol 3-kinase-related kinase.

Pol α. DNA polymerase alpha.

Pol δ. DNA polymerase delta.

Pol ε. DNA polymerase epsilon.

Pol ζ. DNA polymerase zeta.

Pol η. DNA polymerase eta.

Pol κ. DNA polymerase kappa.

PP2A. Protein phosphatase 2A.

PTPA. Protein phosphatase 2 phosphatase activator (PP2A activator).

RIR. Replication-independent ICL repair.

RPA. Replication protein A.

SBS. Single-base substitution.

SC. Supercoiled.

TC-NER. Transcription-coupled nucleotide excision repair.

TLS. Translesion synthesis.

TMEJ. Theta-mediated end joining.

TRAIP. TRAF-interacting protein.

WRN. Werner syndrome RecQ-like helicase.

Chapter 1

INTRODUCTION

1.1 DNA Interstrand Cross-Links: Causes and Consequences

DNA interstrand cross-links (ICLs) are a highly cytotoxic form of DNA damage consisting of a covalent linkage between the two complementary strands of the double helix. ICLs are produced by a chemically diverse set of agents, including endogenous cellular metabolites such as formaldehyde, acetaldehyde, crotonaldehyde, and abasic (AP) sites in their aldehyde forms; microbiome-derived secondary metabolites such as colibactin; and chemotherapeutics such as cisplatin, nitrogen mustard derivatives, and psoralens[1,2]. Although the ICLs produced by these various agents differ significantly in their structure and chemistry, they share the fundamental property of covalently tethering the two DNA strands, thereby preventing the strand unwinding required for DNA replication and transcription. Consequently, ICLs are among the most dangerous lesions that cells encounter, with even a single unrepaired ICL being potentially lethal to a eukaryotic cell[3].

During eukaryotic DNA replication, a replication fork that collides with an ICL stalls because the replicative CDC45-MCM2-7-GINS (CMG) helicase cannot unwind the cross-linked template strands. If the lesion is not resolved, the stalled fork can collapse into a DNA double-strand break (DSB), which can result in large-scale deletions, chromosomal rearrangements, and chromosome missegregation[1,4]. The cytotoxicity of ICLs has been leveraged in the clinic: ICL-inducing agents like cisplatin and mitomycin C are commonly used in chemotherapy, as they can kill rapidly proliferating tumor cells by overwhelming their ICL repair capacity[5]. Conversely, mutations that impair ICL repair give rise to serious

human diseases, most prominently Fanconi anemia (FA), a rare inherited syndrome characterized by progressive bone marrow failure, developmental abnormalities, and a predisposition to various cancers, including acute myeloid leukemia and squamous cell carcinomas[6]. The sources of endogenous ICLs that drive FA pathology have been the subject of considerable investigation. Genetic studies in mice have demonstrated synergy between gene mutations associated with FA and mutations in *ALDH2* and *ADH5*, which are responsible for the processing and clearance of endogenous aldehydes, suggesting that aldehydes produced during normal metabolism are likely to be a physiologically relevant source of ICL damage[7-9]. More recently, attention has turned to microbiome-derived genotoxins as another source of endogenous ICLs. Colibactin, a secondary metabolite produced by *pks*+ *Escherichia coli* strains harbored in the human gut, has emerged as a clinically important ICL-inducing agent, and it will be discussed at length in Chapter 2[10]. Given the relevance of ICL repair to cancer therapy, genetic diseases, and genome maintenance in response to endogenous genotoxins, it is unsurprising that the DNA damage response field has devoted considerable effort to characterizing a variety of ICL repair mechanisms.

1.2 Mechanisms of ICL Repair

To contend with the cytotoxicity of ICLs, cells have evolved multiple mechanisms to resolve these lesions. These mechanisms can be broadly divided into replication-dependent and replication-independent pathways, though the relative contributions of each depend on cell type, cell cycle phase, transcriptional activity at the ICL locus, and the structural properties of the specific ICL[1].

Replication-independent ICL repair (RIR) can process ICLs without the involvement of a replication fork, allowing for ICL resolution to occur in the G1 phase of the cell cycle[11]. In *Xenopus* egg extracts that do not support DNA replication or transcription, certain ICLs can be recognized and processed by mismatch repair (MMR) machinery: the MutSa complex (MSH2-MSH6) acts as an ICL sensor, detecting helical distortion caused by the cross-link and recruiting MutLa and EXO1 to initiate repair[12]. In mammalian cells and extracts, RIR can proceed by global genome nucleotide excision repair (GG-NER), particularly when the ICL is helix-distorting and can be recognized by the NER damage sensor XPC, or by transcription-coupled nucleotide excision repair (TC-NER)[11,13,14]. Despite these insights, the physiological significance of RIR remains incompletely understood. While the mild sensitivity of NER-deficient vertebrate cells to ICL-inducing agents has been taken to suggest that RIR plays a relatively minor role, this interpretation may underestimate its true contribution – since unrepaired ICLs are considerably more toxic in S phase than in G1, replication-coupled repair dominates the measurable cellular response to ICL-inducing agents, potentially obscuring the importance of RIR in non-dividing or slowly cycling cells[1,11]. Nevertheless, replication-coupled ICL repair is generally considered the primary mechanism for cellular resistance to ICL-inducing agents in proliferating cells and accordingly constitutes the focus of this thesis.

During replication-coupled ICL repair, the encounter between a replication fork and an ICL can trigger a number of different repair mechanisms. In mammalian cells, a single replication fork can bypass an ICL through a mechanism called ICL traverse, in which the replisome translocates past the cross-link without resolving it, reinitiating DNA synthesis downstream of the lesion[15]. Traverse seems to account for the majority of encounters

between replication machinery and ICLs in human cells, and it has been shown to require the translocase FANCM, the BLM-TOP3A-RMI1-2 (BTR) complex, and the primase-polymerase PrimPol, which performs the repriming events necessary to restart DNA synthesis downstream of the ICL[15-17]. The molecular mechanisms by which the replisome traverses the cross-linked template and by which the cross-link is resolved following traverse have not yet been fully elucidated[1]. If two replication forks converge at an ICL, the first line of defense for certain ICLs is direct unhooking: acetaldehyde-induced ICLs can be unhooked by an as yet unidentified enzyme, while psoralen-ICLs and AP-ICLs can be unhooked by the DNA glycosylase NEIL3, which cleaves one of the *N*-glycosyl bonds in the cross-link[18,19]. Notably, neither of these pathways involves generating a DSB intermediate, preventing the chromosomal instability that can accompany DSB generation and repair. Should these ICL structure-specific pathways fail, the ICL is processed by the Fanconi anemia pathway, which is the most broadly active and best-characterized mechanism for replication-coupled ICL repair and whose disruption underlies the pathogenesis of Fanconi anemia[1].

1.3 The Fanconi Anemia Pathway

As stated above, Fanconi anemia is characterized by hypersensitivity to ICL-inducing agents and by characteristic chromosomal instability in response to such agents – features that led to the recognition that FA results from a defect in ICL repair. To date, mutations in any of 23 *FANC* genes, whose corresponding proteins function together in the FA pathway to coordinate ICL repair, have been linked to Fanconi anemia[1]. Importantly, several FA pathway proteins, including BRCA1 (FANCS), BRCA2 (FANCD1), PALB2

(FANCN), and RAD51C (FANCO), are also major breast and ovarian cancer susceptibility genes, underscoring the broader clinical relevance of the pathway.

The molecular mechanism of FA pathway-mediated ICL repair has been elucidated through a combination of biochemical studies in *Xenopus* egg extracts and genetic and cell biological approaches in mammalian cells. The pathway is activated when two replication forks converge and stall at an ICL, leading to polyubiquitylation of both stalled CMG helicases by the replisome-associated E3 ubiquitin ligase TRAIP[4,20]. TRAIP also regulates pathway choice: shorter ubiquitin chains on CMG recruit NEIL3 via its ubiquitin-binding domain, allowing it to attempt ICL unhooking; if it fails, TRAIP can continue to extend the ubiquitin chains until polyubiquitylated CMG is unloaded from chromatin by the p97 AAA+ ATPase[20,21]. CMG unloading allows the leading strand to advance from its stall position ~20-40 nucleotides from the ICL to within one nucleotide of the cross-link, at which point one of the replication forks undergoes reversal: the parental strands reanneal and the displaced nascent strands anneal to each other, forming a four-helix junction[22]. Although a requirement for fork reversal during ICL repair has not been formally demonstrated, it appears to represent an on-pathway intermediate[1]. Notably, the four-helix junction generated by fork reversal positions the ICL within a DNA duplex at a three-helix junction – the preferred substrate for the XPF-ERCC1 3' flap endonuclease responsible for unhooking the ICL – providing a structural rationale for why fork reversal may facilitate ICL resolution[23]. The FA core complex, an E3 ubiquitin ligase composed of eight FANC proteins, then monoubiquitylates the FANCI-FANCD2 heterodimer, stabilizing its clamp-like structure on the DNA and allowing it to recruit structure-specific endonucleases to perform ICL unhooking[24,25]. XPF-ERCC1, in complex with the scaffold protein SLX4,

makes incisions on either side of the ICL on one parental strand, leaving one sister chromatid with an ICL remnant (a monoadduct covalently attached to the DNA strand) and the other with a double-strand break[26]. The precise mechanism of the second incision – whether it is performed by XPF-ERCC1 alone or with the assistance of SNM1A or another nuclease – remains an open question in the field[27,28]. After the reversed fork is restored back to a three-way junction, the ICL remnant is bypassed by the translesion synthesis (TLS) polymerase complex Pol κ-REV1-Pol ζ, while the DSB is repaired through RAD51-dependent homologous recombination using the intact sister chromatid as a template[29-31].

A critical and long-underappreciated step in the FA pathway is the remodeling of the stalled replication fork following ICL unhooking. The fork reversal that occurs prior to unhooking can involve extensive reannealing of parental and nascent strands, so the process of "restoring" the fork – returning the 3' end of the nascent strand to the unhooked ICL remnant in order to initiate TLS – is non-trivial. While factors such as the RECQ1 helicase and the DNA2 helicase/nuclease have been implicated in the restoration of reversed forks, the mechanism by which fork restoration is regulated during ICL repair by the FA pathway has remained poorly understood[32,33]. This open question is a central focus of Chapter 3 of this thesis.

1.4 Colibactin: A Microbiome-Derived ICL-Inducing Genotoxin

Chapter 2 of this thesis deals with the repair of ICLs induced by the microbial genotoxin colibactin. The composition and activities of the gut microbiome have an immense impact on human health, and emerging evidence shows that pathogenic gut bacteria can damage host DNA by producing genotoxic secondary metabolites[34]. Colibactin, produced by *pks*+ *Escherichia coli* strains harboring a 52-kilobase biosynthetic gene cluster, has attracted

intense interest as a driver of human disease, particularly colorectal cancer[35]. Indeed, colibactin-producing bacteria are detected in approximately 40% of patients with inflammatory bowel disease and 60% of patients with colorectal cancer[36]. Notably, colibactin exposure is associated with specific mutational signatures – single-base substitutions and insertions/deletions at AT-rich sequences – present in the colonic crypts of not only cancer patients, but also most healthy individuals, implicating colibactin in the mutagenic landscape of the colon[37-39].

Colibactin is a pseudo-C_2-symmetric molecule bearing two cyclopropane "warheads," each of which can react with a deoxyadenosine residue[10,40]. Reaction of one of these cyclopropanes with a deoxyadenosine forms a colibactin monoadduct; reaction of both cyclopropanes with deoxyadenosines on opposite strands of host cell DNA produces a DNA-ICL that covalently links the two strands at offset adenines[10,41,42]. The colibactin-ICL (clb-ICL) makes the cross-linked adenosines particularly susceptible to depurination, which leaves one strand with an N3 monoalkylated base and the other with an AP site, which could subsequently be converted into a single-strand break[43]. Thus, in addition to ICLs, colibactin exposure produces a wide spectrum of other DNA lesions, including DSBs and various monoalkylation adducts, reflecting the complex chemistry of colibactin's bifunctional warheads. Despite the chemical characterization of the colibactin-ICL and the epidemiological links between colibactin exposure and colorectal cancer, the cellular mechanisms by which colibactin-induced ICLs are repaired had not been clarified prior to the work described in Chapter 2.

1.5 The *Xenopus* Egg Extract System

Much of what is known about the molecular mechanism of ICL repair in vertebrates has been learned through the use of cell-free extracts prepared from the eggs of the African clawed frog, *Xenopus laevis*. This experimental system, first developed for studying DNA replication in the 1980's and subsequently adapted by the laboratory of Johannes Walter to study site-specific DNA lesions, has proven uniquely powerful for investigating replication-coupled ICL repair at biochemical resolution[3,44,45]. Crushing *Xenopus* eggs by centrifugation releases their cytoplasmic contents, which contain abundant stores of replication and repair factors stockpiled in preparation for the rapid embryonic cell divisions that follow fertilization[45]. The presence of these factors allows the extracts to faithfully recapitulate the key events of vertebrate DNA replication, including licensing, initiation, elongation, and termination, as well as the events of multiple DNA repair pathways.

Our system uses two sequential extracts: a G1 phase high-speed supernatant (HSS), which chromatinizes added DNA and loads the MCM2-7 replicative helicase onto origins of replication, and an S phase nucleoplasmic extract (NPE), which contains high concentrations of CDK2-cyclin E and initiation factors that trigger a single, synchronous round of DNA replication[45,46]. Because replication can be initiated on thousands of exogenously added plasmid molecules simultaneously, the system enables temporal ordering of replication and repair events, allowing observation even at single-nucleotide resolution. The extracts are also highly amenable to biochemical manipulation: specific proteins can be quantitatively immunodepleted and replaced with recombinant wild-type or mutant variants, providing a powerful depletion and rescue framework that is not readily achievable in cell-based systems. Furthermore, when plasmid backbones are engineered to

contain a *lacO* array, chromatin-associated proteins can be recovered by plasmid pulldown using magnetic beads coated with the lac repressor (LacI), allowing for Western blot and mass spectrometry analysis[29,47].

Critically for the study of ICL biology, the Walter laboratory and its collaborators developed methods of incorporating site-specific DNA lesions into plasmid templates at defined positions[3]. These precisely engineered plasmids combined with the synchronous, biochemically tractable *Xenopus* egg extract system allow the molecular events at a single ICL to be followed in real time. This approach was used to establish the first cell-free system for replication-coupled ICL repair and to define the basic sequence of events: fork stalling, leading strand approach up to the cross-link, ICL unhooking, TLS, and DSB repair by homologous recombination[3]. Subsequent work from multiple laboratories has used the egg extract system to determine the requirements for CMG ubiquitylation and unloading[20,21], fork reversal[22], FANCI-FANCD2 ubiquitylation and function[48], ICL unhooking by XPF-ERCC1[26], TLS[3,29], and HR-mediated DSB repair[31]. The egg extract system has thus provided much of the mechanistic framework for our current understanding of FA pathway-mediated ICL repair.

Beyond their utility as a reconstituted repair system, *Xenopus* egg extracts offer several advantages over cell-based approaches that are particularly relevant to the work described in this thesis. Firstly, because DNA replication is synchronous and plasmid-based, repair intermediates can be captured and analyzed at defined time points, and nascent strand extension can be observed with single-nucleotide resolution. Secondly, the absence of a nucleus means that all soluble proteins, including the factors necessary to recapitulate replication and repair, are present in a single biochemical compartment and can be

manipulated directly (for instance, through depletion or inhibition with small molecules). Finally, the plasmid pulldown approach enables quantitative proteomics of repair factors at defined DNA lesions at specific times after replication initiation, providing a resolution that is difficult to achieve in cell-based systems. These features have made the *Xenopus* egg extract system the preeminent tool for dissecting the molecular mechanism of ICL repair in vertebrates, and accordingly, the majority of the experimental work described in this thesis was performed using this system.

1.6 Open Questions

Despite the substantial progress outlined above, two major gaps in understanding ICL repair motivated the work described in this thesis. The first concerns the repair of ICLs induced by colibactin. As described above, colibactin is strongly implicated as a driver of colorectal carcinogenesis, and the clb-ICL represents a naturally occurring, physiologically relevant lesion whose formation has been confirmed in living cells, making its repair mechanism an important open question both clinically and as the first opportunity to examine FA pathway activity on an endogenous ICL substrate. Although the chemistry of colibactin and the lesions it produces has been studied in detail, the cellular mechanisms by which colibactin-induced ICLs are repaired had not been elucidated. Prior to the work described in this thesis, there was indirect evidence of FA pathway-mediated repair of colibactin-ICLs: colibactin exposure was shown to induce FANCD2 monoubiquitylation and foci formation in cells, and FANCD2-deficient cells exhibited hypersensitivity to colibactin[49]. However, these observations could reflect the cellular response to a variety of colibactin-induced lesions rather than ICLs specifically, as colibactin produces a spectrum of DNA damage and FANCD2 participates in multiple DNA protection and repair

pathways. Therefore, it was unclear whether the FA pathway or an alternative mechanism is responsible for processing clb-ICLs, and this question is the focus of Chapter 2.

The second open question concerns the regulation of replication fork dynamics during FA pathway-mediated ICL repair. As described above, fork reversal and subsequent fork restoration are critical steps in the pathway, but the regulatory mechanisms that control these fork remodeling events are poorly understood. Members of the phosphatidylinositol 3-kinase-related kinase (PIKK) family were known to play roles in coordinating fork remodeling: in particular, DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and the ataxia-telangiectasia mutated (ATM) kinase stimulate fork reversal under replication stress conditions[50-52], and ATM and the ataxia telangiectasia and Rad3-related (ATR) kinases promote replication fork protection and recovery by phosphorylating resection enzymes, such as the helicase WRN and the nuclease MRE11[53,54]. However, how these kinases themselves are activated and regulated specifically in response to fork dynamics at ICLs, which nucleases they coordinate to enable fork restoration, and how the extent of resection is controlled to prevent aberrant processing and fork collapse remained open questions. These questions are the focus of Chapter 3, which identifies ATM as the kinase responsible for activating reversed fork resection during ICL repair, defines the nuclease activities it coordinates, and characterizes the opposing phosphatase activity that constrains resection to ensure productive fork restoration.

1.7 Summary of This Thesis

Chapter 2 of this thesis describes the elucidation of the mechanism of repair of colibactin-induced DNA-ICLs. Using *Xenopus* egg extracts, we demonstrate that replication fork stalling at a clb-ICL activates the FA pathway, which resolves the cross-link through

nucleolytic incisions that generate a DSB intermediate in one sister chromatid and an ICL remnant in the other. We characterize the translesion synthesis polymerases responsible for bypassing the clb-ICL remnant, finding that Pol η and the Pol κ-REV1-Pol ζ complex act sequentially at the site of colibactin alkylation. Finally, we show that TLS past the clb-ICL remnant introduces predominantly T>A point mutations, providing a molecular framework for understanding how colibactin exposure drives mutagenesis in the gut epithelium.

Chapter 3 describes the characterization of an ATM kinase-dependent regulatory axis that controls nucleolytic processing of reversed fork intermediates during FA pathway-mediated ICL repair. We find that ATM is activated concomitantly with fork reversal and stimulates resection of reversed forks through the coordinated activities of the EXO1 and DNA2 nucleases, with EXO1 initiating resection of nascent lagging strands in the regressed arm and DNA2 performing subsequent processing of recessed lagging strands. We further show that ATM-dependent resection is opposed by activity of the serine/threonine phosphatase PP2A, revealing a phospho-regulatory axis that tunes the extent of fork resection to ensure productive fork restoration and enable subsequent bypass of the ICL remnant by TLS polymerases. Together, these findings establish ATM as a central coordinator of fork remodeling during ICL repair and uncover a kinase-phosphatase axis that safeguards genome integrity during this process.

Chapter 4 presents preliminary data investigating the mechanism by which replication fork structures generated during ICL repair activate the FA pathway independently of ongoing DNA replication. Using deproteinized, gel-purified replication intermediates added to egg extract that does not support replication, we show that X-shaped structures are sufficient to activate ATR signaling and FANCD2 monoubiquitylation, both of which are consistent

with FA pathway activation. We further demonstrate that the X-structures can be resolved into repaired monomeric products in the absence of replication, suggesting that the DNA structures generated by fork convergence, rather than the replisome itself, may trigger FA pathway-mediated processing. These findings lay the groundwork for future investigations into how the FA pathway is activated and which specific structural features of replication intermediates are recognized to initiate ICL unhooking.

Together, these studies advance our mechanistic understanding of how cells tolerate one of the most cytotoxic forms of DNA damage and illuminate regulatory mechanisms that ensure accurate and efficient ICL repair.

1.8 References

1 Semlow, D. R. & Walter, J. C. Mechanisms of Vertebrate DNA Interstrand Cross-Link Repair. *Annu Rev Biochem* **90**, 107-135 (2021). https://doi.org/10.1146/annurev-biochem-080320-112510

2 Deans, A. J. & West, S. C. DNA interstrand crosslink repair and cancer. *Nat Rev Cancer* **11**, 467-480 (2011). https://doi.org/10.1038/nrc3088

3 Raschle, M. *et al.* Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* **134**, 969-980 (2008). https://doi.org/10.1016/j.cell.2008.08.030

4 Zhang, J. *et al.* DNA interstrand cross-link repair requires replication-fork convergence. *Nat Struct Mol Biol* **22**, 242-247 (2015). https://doi.org/10.1038/nsmb.2956

5 Duxin, J. P. & Walter, J. C. What is the DNA repair defect underlying Fanconi anemia? *Curr Opin Cell Biol* **37**, 49-60 (2015). https://doi.org/10.1016/j.ceb.2015.09.002

6 Niraj, J., Farkkila, A. & D'Andrea, A. D. The Fanconi Anemia Pathway in Cancer. *Annu Rev Cancer Biol* **3**, 457-478 (2019). https://doi.org/10.1146/annurev-cancerbio-030617-050422

7 Langevin, F., Crossan, G. P., Rosado, I. V., Arends, M. J. & Patel, K. J. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* **475**, 53-58 (2011). https://doi.org/10.1038/nature10192

8 Rosado, I. V., Langevin, F., Crossan, G. P., Takata, M. & Patel, K. J. Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. *Nat Struct Mol Biol* **18**, 1432-1434 (2011). https://doi.org/10.1038/nsmb.2173

9 Pontel, L. B. *et al.* Endogenous Formaldehyde Is a Hematopoietic Stem Cell Genotoxin and Metabolic Carcinogen. *Mol Cell* **60**, 177-188 (2015). https://doi.org/10.1016/j.molcel.2015.08.020

10 Wilson, M. R. *et al.* The human gut bacterial genotoxin colibactin alkylates DNA. *Science* **363** (2019). https://doi.org/10.1126/science.aar7785

11 Enoiu, M., Jiricny, J. & Scharer, O. D. Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic Acids Res* **40**, 8953-8964 (2012). https://doi.org/10.1093/nar/gks670

12 Kato, N. *et al.* Sensing and Processing of DNA Interstrand Crosslinks by the Mismatch Repair Pathway. *Cell Rep* **21**, 1375-1385 (2017). https://doi.org/10.1016/j.celrep.2017.10.032

13 Muniandy, P. A., Thapa, D., Thazhathveetil, A. K., Liu, S. T. & Seidman, M. M. Repair of laser-localized DNA interstrand cross-links in G1 phase mammalian cells. *J Biol Chem* **284**, 27908-27917 (2009). https://doi.org/10.1074/jbc.M109.029025

14 Smeaton, M. B. *et al.* Distortion-dependent unhooking of interstrand cross-links in mammalian cell extracts. *Biochemistry* **47**, 9920-9930 (2008). https://doi.org/10.1021/bi800925e

15 Huang, J. *et al.* The DNA translocase FANCM/MHF promotes replication traverse of DNA interstrand crosslinks. *Mol Cell* **52**, 434-446 (2013). https://doi.org/10.1016/j.molcel.2013.09.021

16 Ling, C. *et al.* Bloom syndrome complex promotes FANCM recruitment to stalled replication forks and facilitates both repair and traverse of DNA interstrand crosslinks. *Cell Discov* **2**, 16047 (2016). https://doi.org/10.1038/celldisc.2016.47

17 Gonzalez-Acosta, D. *et al.* PrimPol-mediated repriming facilitates replication traverse of DNA interstrand crosslinks. *EMBO J* **40**, e106355 (2021). https://doi.org/10.15252/embj.2020106355

18 Hodskinson, M. R. *et al.* Alcohol-derived DNA crosslinks are repaired by two distinct mechanisms. *Nature* **579**, 603-608 (2020). https://doi.org/10.1038/s41586-020-2059-5

19 Semlow, D. R., Zhang, J., Budzowska, M., Drohat, A. C. & Walter, J. C. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase. *Cell* **167**, 498-511 e414 (2016). https://doi.org/10.1016/j.cell.2016.09.008

20 Wu, R. A. *et al.* TRAIP is a master regulator of DNA interstrand crosslink repair. *Nature* **567**, 267-272 (2019). https://doi.org/10.1038/s41586-019-1002-0

21 Fullbright, G., Rycenga, H. B., Gruber, J. D. & Long, D. T. p97 Promotes a Conserved Mechanism of Helicase Unloading during DNA Cross-Link Repair. *Mol Cell Biol* **36**, 2983-2994 (2016). https://doi.org/10.1128/MCB.00434-16

22 Amunugama, R. *et al.* Replication Fork Reversal during DNA Interstrand Crosslink Repair Requires CMG Unloading. *Cell Rep* **23**, 3419-3428 (2018). https://doi.org/10.1016/j.celrep.2018.05.061

23 Abdullah, U. B. *et al.* RPA activates the XPF-ERCC1 endonuclease to initiate processing of DNA interstrand crosslinks. *EMBO J* **36**, 2047-2060 (2017). https://doi.org/10.15252/embj.201796664

24 Garcia-Higuera, I. *et al.* Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol Cell* **7**, 249-262 (2001). https://doi.org/10.1016/s1097-2765(01)00173-3

25 Alcón, P. *et al.* FANCD2-FANCI is a clamp stabilized on DNA by monoubiquitination of FANCD2 during DNA repair. *Nat Struct Mol Biol* **27**, 240-248 (2020). https://doi.org/10.1038/s41594-020-0380-1

26 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). https://doi.org/10.1016/j.molcel.2014.03.015

27 Wang, A. T. *et al.* Human SNM1A and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair. *Genes Dev* **25**, 1859-1870 (2011). https://doi.org/10.1101/gad.15699211

28 Hoogenboom, W. S., Boonen, R. & Knipscheer, P. The role of SLX4 and its associated nucleases in DNA interstrand crosslink repair. *Nucleic Acids Res* **47**, 2377-2388 (2019). https://doi.org/10.1093/nar/gky1276

29 Budzowska, M., Graham, T. G., Sobeck, A., Waga, S. & Walter, J. C. Regulation of the Rev1-pol ζ complex during bypass of a DNA interstrand cross-link. *EMBO J* **34**, 1971-1985 (2015). https://doi.org/10.15252/embj.201490878

30 Sellés-Baiget, S. *et al.* Catalytic and noncatalytic functions of DNA polymerase κ in translesion DNA synthesis. *Nature Structural & Molecular Biology* (2024). https://doi.org/10.1038/s41594-024-01395-3

31 Long, D. T., Räschle, M., Joukov, V. & Walter, J. C. Mechanism of RAD51-dependent DNA interstrand cross-link repair. *Science* **333**, 84-87 (2011). https://doi.org/10.1126/science.1204258

32 Berti, M. *et al.* Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition. *Nat Struct Mol Biol* **20**, 347-354 (2013). https://doi.org/10.1038/nsmb.2501

33 Thangavel, S. *et al.* DNA2 drives processing and restart of reversed replication forks in human cells. *J Cell Biol* **208**, 545-562 (2015). https://doi.org/10.1083/jcb.201406100

34 Cao, Y. *et al.* Commensal microbiota from patients with inflammatory bowel disease produce genotoxic metabolites. *Science* **378**, eabm3233 (2022). https://doi.org/10.1126/science.abm3233

35 Wernke, K. M. *et al.* Structure and bioactivity of colibactin. *Bioorg Med Chem Lett* **30**, 127280 (2020). https://doi.org/10.1016/j.bmcl.2020.127280

36 Arthur, J. C. *et al.* Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* **338**, 120-123 (2012). https://doi.org/10.1126/science.1224820

37 Dziubańska-Kusibab, P. J. *et al.* Colibactin DNA-damage signature indicates mutational impact in colorectal cancer. *Nat Med* **26**, 1063-1069 (2020). https://doi.org/10.1038/s41591-020-0908-2

38 Pleguezuelos-Manzano, C. *et al.* Mutational signature in colorectal cancer caused by genotoxic pks. *Nature* **580**, 269-273 (2020). https://doi.org/10.1038/s41586-020-2080-8

39 Lee-Six, H. *et al.* The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* **574**, 532-537 (2019). https://doi.org/10.1038/s41586-019-1672-7

40 Xue, M. *et al.* Structure elucidation of colibactin and its DNA cross-links. *Science* **365** (2019). https://doi.org/10.1126/science.aax2685

41 Jiang, Y. *et al.* Reactivity of an Unusual Amidase May Explain Colibactin's DNA Cross-Linking Activity. *J Am Chem Soc* **141**, 11489-11496 (2019). https://doi.org/10.1021/jacs.9b02453

42 Carlson, E. S. *et al.* The specificity and structure of DNA cross-linking by the gut bacterial genotoxin colibactin. *Science* **390**, eady3571 (2025). https://doi.org/10.1126/science.ady3571

43 Xue, M., Wernke, K. M. & Herzon, S. B. Depurination of Colibactin-Derived Interstrand Cross-Links. *Biochemistry* **59**, 892-900 (2020). https://doi.org/10.1021/acs.biochem.9b01070

44 Blow, J. J. & Laskey, R. A. Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of Xenopus eggs. *Cell* **47**, 577-587 (1986). https://doi.org/10.1016/0092-8674(86)90622-7

45 Walter, J., Sun, L. & Newport, J. Regulated chromosomal DNA replication in the absence of a nucleus. *Mol Cell* **1**, 519-529 (1998). https://doi.org/10.1016/s1097-2765(00)80052-0

46 Lebofsky, R., Takahashi, T., and Walter, J.C. in *Methods in Molecular Biology, DNA Replication* Vol. 521 (ed Sonya Vengrova and Jacob Z. Dalgaard) Ch. 13, 229-252 (Humana Press, 2009).

47 Räschle, M. *et al.* DNA repair. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. *Science* **348**, 1253671 (2015). https://doi.org/10.1126/science.1253671

48 Knipscheer, P. *et al.* The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* **326**, 1698-1701 (2009). https://doi.org/10.1126/science.1182372

49 Bossuet-Greif, N. *et al.* The Colibactin Genotoxin Generates DNA Interstrand Cross-Links in Infected Cells. *mBio* **9** (2018). https://doi.org/10.1128/mBio.02393-17

50 Schleicher, E. M. *et al.* The TIP60-ATM axis regulates replication fork stability in BRCA-deficient cells. *Oncogenesis* **11**, 33 (2022). https://doi.org/10.1038/s41389-022-00410-w

51 Bansbach, C. E., Betous, R., Lovejoy, C. A., Glick, G. G. & Cortez, D. The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks. *Genes Dev* **23**, 2405-2414 (2009). https://doi.org/10.1101/gad.1839909

52 Dibitetto, D. *et al.* DNA-PKcs promotes fork reversal and chemoresistance. *Mol Cell* **82**, 3932-3942 e3936 (2022). https://doi.org/10.1016/j.molcel.2022.08.028

53 Ammazzalorso, F., Pirzio, L. M., Bignami, M., Franchitto, A. & Pichierri, P. ATR and ATM differently regulate WRN to prevent DSBs at stalled replication forks and promote replication fork recovery. *EMBO J* **29**, 3156-3169 (2010). https://doi.org/10.1038/emboj.2010.205

54 Zhang, H. *et al.* ATM priming and end resection-coupled phosphorylation of MRE11 is important for fork protection and replication restart. *Proc Natl Acad Sci U S A* **122**, e2422720122 (2025). https://doi.org/10.1073/pnas.2422720122

Chapter 2

THE FANCONI ANEMIA PATHWAY REPAIRS COLIBACTIN-INDUCED DNA INTERSTRAND CROSS-LINKS

Maria Altshuller[1], Xu He[1], Elliot J. MacKrell[1], Kevin M. Wernke[2], Yougant Airan[2], Joel W. H. Wong[3], Selene Sellés-Baiget[4], Ting-Yu Wang[5,6], Tsui-Fen Chou[5,6], Julien P. Duxin[4,7], Emily P. Balskus[3,8], Seth B. Herzon[2], and Daniel R. Semlow[1]*

Affiliations

[1]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA

[2]Department of Chemistry, Yale University, New Haven, CT, USA.

[3]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA

[4]The Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[5]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA

[6]Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, Pasadena, CA, USA

[7]Biotech Research and Innovation Centre, Faculty of Health and Medical Sciences, University of Copenhagen, DK-2200, Copenhagen, Denmark

[8]Howard Hughes Medical Institute, Harvard University, Cambridge, MA, USA

*Corresponding author (dsemlow@caltech.edu)

This chapter consists of published work that is available at:

Altshuller, M. *et al.* The Fanconi anemia pathway repairs colibactin-induced DNA interstrand cross-links. *Nat Commun* **16**, 10563 (2025). https://doi.org/10.1038/s41467-025-65606-1

2.1 Abstract

Colibactin is a secondary metabolite produced by bacteria present in the human gut and is implicated in the development of colorectal cancer. This genotoxin alkylates deoxyadenosines on opposite strands of host cell DNA to produce DNA interstrand cross-links. While cells have evolved multiple mechanisms to resolve ("unhook") interstrand cross-links, little is known about which of these pathways promote resistance to colibactin. Here, we use *Xenopus* egg extracts to investigate replication-coupled repair of colibactin-induced interstrand cross-links. We show that replication fork stalling at a colibactin-induced interstrand cross-link activates the Fanconi anemia interstrand cross-link repair pathway, which unhooks the interstrand cross-link through nucleolytic incisions. These incisions generate a DNA double-strand break intermediate in one sister chromatid, which can be repaired by homologous recombination, and a monoadduct ("interstrand cross-link remnant") in the other. Translesion synthesis past the colibactin-induced interstrand cross-link remnant depends on Pol η and the Pol κ-REV1-Pol ζ polymerase complex and introduces predominantly T>A point mutations at the sites of colibactin alkylation. Taken together, our work provides a molecular framework for understanding how cells tolerate a naturally occurring and clinically relevant interstrand cross-link.

2.2 Introduction

Human health is profoundly influenced by host-microbiome interactions. Emerging data indicate that pathogenic gut bacteria can contribute to dysbiosis and promote inflammation by producing toxins that damage host DNA[1]. *pks*$^+$ bacteria harbor a 52 kilobase (kb) non-ribosomal polypeptide synthetase/polyketide synthase biosynthetic gene cluster that enables production of the colibactin genotoxin[2]. Colibactin-producing bacteria are detected in ~40% of patients with inflammatory bowel disease (IBD) and ~60% of patients with colorectal cancer[3]. Moreover, colibactin exposure is associated with single-base substitution (SBS) and insertion/deletion (ID) mutational signatures found in the colonic crypts of most healthy individuals[4-6]. Colibactin is therefore a highly prevalent and potent driver of human disease.

Colibactin (clb) comprises a pseudo-C_2-symmetric scaffold with two cyclopropane groups (Fig. 2.1a)[7-9]. Each cyclopropane can react with a deoxyadenosine (dA) to form a monoadduct or a DNA interstrand cross-link (ICL) that covalently links the two DNA strands. Consistently, clb-associated mutational signatures target AT-rich motifs containing offset dAs on opposite strands[5,6,10]. Once formed, the colibactin-ICL (clb-ICL) renders the alkylated dAs susceptible to depurination, resulting in an N3 monoalkylated base in one strand and a DNA abasic (AP) site in the other. The AP site can further decompose into a DNA strand break[11]. Shunt colibactin metabolites that have been isolated from *pks*$^+$ mutated strains may also alkylate DNA[12]. Thus, *pks*$^+$ bacteria likely generate a complicated spectrum of host DNA lesions that mobilizes multiple DNA repair pathways.

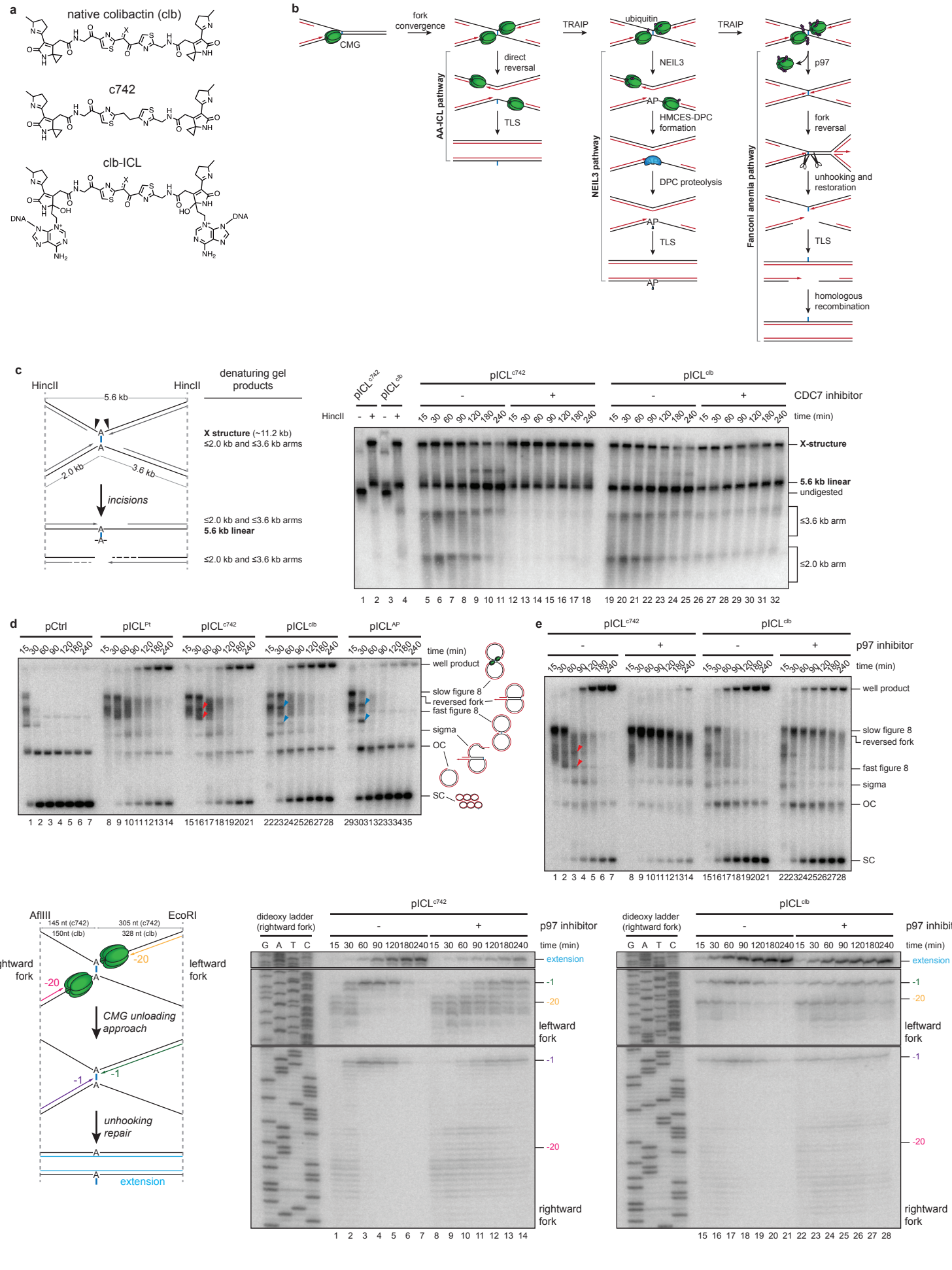

a
native colibactin (clb)
c742
clb-ICL
DNA
b
CMG
fork convergence
direct reversal
TLS
AA-ICL pathway
TRAIP
ubiquitin
NEIL3
AP
HMCES-DPC formation
DPC proteolysis
NEIL3 pathway
p97
fork reversal
unhooking and restoration
Fanconi anemia pathway
homologous recombination
c
HincII
denaturing gel products
5.6 kb
2.0 kb
3.6 kb
incisions
X structure (~11.2 kb)
≤2.0 kb and ≤3.6 kb arms
5.6 kb linear
pICLc742
pICLclb
CDC7 inhibitor
time (min)
X-structure
5.6 kb linear
undigested
≤3.6 kb arm
≤2.0 kb arm
d
pCtrl
pICLPt
pICLAP
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
e
p97 inhibitor
f
AflIII
EcoRI
145 nt (c742)
150nt (clb)
305 nt (c742)
328 nt (clb)
rightward fork
leftward fork
CMG unloading approach
unhooking repair
extension
dideoxy ladder (rightward fork)
G A T C
-1
-20

Fig. 2.1 | Replication-coupled repair of colibactin-induced ICLs

a, Structures of the native colibactin toxin (top; X indicates O or NH_2), stabilized c742 colibactin analog (middle), and the colibactin-induced ICL (clb-ICL; bottom).

b, Schematic of replication-coupled ICL repair pathways (see main text for details). Black, parental DNA strands. Red, nascent DNA strands. Blue, ICL. Purple, ubiquitin.

c, Left, schematic of species generated during ICL repair. Digestion with HincII produces X-structures corresponding to cross-linked plasmids and linear species corresponding to unhooked plasmids. Right, $pICL^{c742}$ and $pICL^{clb}$ were replicated in egg extract supplemented with CDC7 inhibitor, as indicated. Recovered replication intermediates were digested with HincII and resolved on a denaturing (alkaline) agarose gel. Intermediates were detected by Southern blotting with ^{32}P radiolabeled probes and visualized by autoradiography.

d, The indicated plasmids were replicated in egg extract supplemented with [α-^{32}P]dCTP. Replication intermediates were separated on a native agarose gel and visualized by autoradiography. Schematics of replication and repair intermediates that are resolved by native agarose gel electrophoresis are shown. SC, supercoiled; OC, open circular. Red arrowheads indicate unique intermediates formed during replication of $pICL^{c742}$; blue arrowheads indicate OC-OC catenae and OC-SC catenae produced during replication of $pICL^{AP}$ and $pICL^{clb}$.

e, $pICL^{c742}$ and $pICL^{clb}$ were replicated in egg extract supplemented with [α-^{32}P]dCTP and p97i, as indicated. Replication intermediates were analyzed as in **d**. Red arrowheads indicate $pICL^{c742}$-specific repair intermediates.

f, Left, schematic of nascent leading strands generated during ICL repair. AflIII cuts 145 or 150 nt to the left of the c742-ICL or clb-ICL, respectively. EcoRI cuts 305 or 328 nt to the right of the c742-ICL or clb-ICL, respectively. Digestion with AflIII and EcoRI generates characteristic nascent strand -20 to -40 stall, -1 stall, and extension products. Right, DNA recovered from the $pICL^{c742}$ and $pICL^{clb}$ replication reactions shown in **e** was digested with AflIII and EcoRI and resolved on a denaturing polyacrylamide sequencing gel. Top, middle, and bottom panels show sections of the same gels that contain the extension products, leftward fork leading strands, and rightward fork leading strands, respectively.

The clb-ICL is likely to be highly cytotoxic due to its presumed ability to stall replication forks. Failure to repair ICLs encountered during replication can lead to under-replication of the genome and/or gross chromosomal rearrangements. Consequently, cells have evolved multiple repair mechanisms that are activated when a replication fork encounters an ICL (Fig. 2.1b)[13]. An ICL induced by acetaldehyde (AA-ICL) can be resolved ("unhooked") through direct reversal of the cross-link[14]. ICLs formed by psoralen or by AP sites can be unhooked by the NEIL3 DNA glycosylase, which excises a cross-linked base to produce an AP site[15]. Should these pathways fail, the ICL is unhooked by the Fanconi anemia (FA) pathway[13]. This pathway is activated when replicative CDC45-MCM2-7-GINS (CMG) helicases stall at an ICL[16]. CMGs are then polyubiquitylated by the replisome-associated TRAIP E3 ubiquitin ligase, leading to p97-dependent CMG unloading and reversal of one of the two converged forks[17-19]. At this stage, the FA core complex monoubiquitylates the FANCI-FANCD2 heterodimer, which becomes clamped onto DNA near the ICL[20,21]. Monoubiquitylated FANCI-FANCD2 recruits structure-specific nucleases including XPF-ERCC1 that unhook the ICL through nucleolytic incisions flanking the ICL[22]. Following restoration of the reversed fork, one sister chromatid contains an ICL remnant that can be bypassed by the REV1-Pol ζ translesion synthesis (TLS) polymerase, while the other chromatid contains a DNA double strand break (DSB) that is repaired by homologous recombination[23,24]. Importantly, the FA pathway is thought to be capable of unhooking any ICL, regardless of cross-link structure, and is therefore regarded as the most important pathway for cellular resistance to ICL-inducing agents. Consistent with this inference, mutations in any of 22 genes associated with this pathway cause a human bone marrow failure and cancer predisposition syndrome.

Despite the importance of the FA pathway for human health, the physiological ICLs targeted by this pathway remain obscure. Synergy between mutations in the FA pathway and mutations in aldehyde metabolism imply that the FA pathway resolves ICLs formed by endogenous aldehydes, such as formaldehyde or acetaldehyde[25,26]. Indeed, a chemically synthesized acetaldehyde-ICL is processed by the FA pathway in *Xenopus* egg extracts[14]. However, ICLs induced by acetaldehyde or any other endogenous metabolite have not been confirmed *in vivo*. Furthermore, FANC proteins have been shown to promote a variety of genome maintenance mechanisms in addition to ICL repair, including homologous recombination, replication fork protection, R-loop resolution, retrotransposon suppression, and mitochondrial genome maintenance[27-31]. Intriguingly, indirect evidence suggests that colibactin-induced ICLs might be repaired by the FA pathway. Colibactin exposure induces FANCD2 ubiquitylation and FANCD2 foci formation in cells, and FANCD2-deficient cells are hypersensitive to colibactin[32]. However, given the spectrum of DNA lesions produced during colibactin exposure and the multiple genoprotective functions of FANCD2, it remains unclear whether colibactin-induced ICLs are repaired by the FA pathway.

Here, we use *Xenopus* egg extracts to investigate replication-coupled repair of site-specific ICLs formed by either the native colibactin toxin or a synthetic colibactin analog. We find that both ICLs are strong impediments to DNA replication that trigger replication-coupled repair by the FA pathway. TLS past both an unhooked colibactin-induced ICL remnant and a colibactin monoadduct requires REV1 and Pol κ. We further show that colibactin-induced ICL repair can be error-free but frequently introduces primarily T>A point mutations. These results define the mechanisms that allow cells to tolerate a DNA lesion that may

drive the development of CRC and demonstrate the FA pathway's capacity to repair physiologically relevant ICLs.

2.3 Results

2.3.1 Colibactin-induced ICL repair is coupled to DNA replication

To investigate colibactin-induced ICL repair in *Xenopus* egg extracts, we prepared plasmids containing site-specific ICLs generated with either native colibactin (clb) or a synthetic colibactin analog (c742; Fig. 2.1a). c742 contains an unsubstituted carbon-carbon linker in place of a central diketone that is susceptible to oxidative cleavage. The stabilized analog produces an ICL that is otherwise structurally identical to a clb-ICL and that can induce cell cycle arrest in human cells[33]. Short oligonucleotide duplexes containing a colibactin alkylation motif were incubated with either c742 or with cultured *pks*$^+$ bacteria. In both cases, cross-links were expected to form between dAs at the second position of a palindromic 5'-A**A**TATT-3' target sequence (Supplementary Fig. 2.1a)[5,6]. Following purification, essentially all the duplexes contained a colibactin-induced ICL (Supplementary Fig. 2.1b); these were then ligated into ~5.6 kb plasmids. At least 75% of the purified c742-ICL plasmid (pICL^{c742}) contained an intact ICL (Supplementary Fig. 2.1c). However, only ~50% of the clb-ICL plasmid (pICL^{clb}) contained an intact ICL (Supplementary Fig. 2.1c), despite preparing the plasmids using highly purified cross-linked oligonucleotides (Supplementary Fig. 2.1b). Interestingly, a substantial fraction of the pICL^{clb} plasmids contained AP sites and monoadducts, presumably consisting of clb linked to adenine (clb-A) (Supplementary Fig. 2.1d,e), which indicates that the clb-ICL can undergo depurination during plasmid preparation. Consistent with previous work, the AP

sites in our pICL^{clb} preparations are rapidly removed in egg extracts (Supplementary Fig. 2.1f)[34]. We therefore infer that replication forks encounter either a clb-ICL or a clb-A monoadduct, but not an AP site, during pICL^{clb} replication in egg extracts.

We first asked whether colibactin-induced ICLs are unhooked in egg extracts and whether such unhooking is DNA replication-dependent. pICL^{c742} and pICL^{clb} were replicated in egg extract, recovered, linearized with a restriction enzyme (HincII), resolved on a denaturing agarose gel, and detected by Southern blotting (Fig. 2.1c). Initially, the plasmids were detected as "X-structures" that migrate slowly due to the presence of a cross-link. The initial accumulation of pICL^{clb} X-structures was reduced by the presence of uncross-linked plasmid in our preparation (Supplementary Fig. 2.1c-f). During replication in egg extract, the X-structures are resolved into linear species, indicating that the ICL is unhooked (Fig. 2.1c and Supplementary Fig. 2.1g). Importantly, the conversion of X-structures to linear species was attenuated in extract supplemented with CDC7 inhibitor (PHA-767491), which blocks DNA replication initiation (Fig. 2.1c and Supplementary Fig. 2.1g)[35]. We conclude that c742-ICL and clb-ICL are unhooked by replication-coupled repair in egg extracts.

pICL^{c742} and pICL^{clb} were next replicated in egg extract supplemented with [α-^{32}P]dCTP, which is incorporated into nascent DNA strands. Replication intermediates were then resolved on a native agarose gel and visualized by autoradiography (Fig. 2.1d). For comparison, we also replicated an undamaged plasmid (pCtrl) and plasmids containing either a cisplatin-ICL (pICL^{Pt}) or an AP-ICL (pICL^{AP}). Whereas replicated pCtrl rapidly accumulated as open circular and supercoiled products, each of the ICL-containing plasmids initially accumulated as slowly migrating "figure 8" intermediates that form when CMG helicases converge and stall on either side of the ICL. In the case of pICL^{AP}, the slow

figure 8s were rapidly converted into open circular and supercoiled products, indicative of unhooking by the NEIL3 glycosylase[15]. In the case of pICL^{c742} and pICL^{clb}, the slow figure 8s were first converted into a series of species with intermediate mobilities and later into open circular and supercoiled products. Note that pICL^{clb} replication resulted in substantial accumulation of open circular and supercoiled plasmids at even early time points, consistent with the presence of uncross-linked plasmid in this preparation. Interestingly, pICL^{c742} and pICL^{clb} replication produced intermediates that closely resembled those formed during replication of pICL^{Pt}, which is repaired by the FA pathway in egg extracts[36,37]. These findings indicate that colibactin-induced ICLs block replication fork progression and activate repair.

In egg extracts, the convergence of replication forks at the ICL triggers TRAIP-dependent polyubiquitylation of CMG[17]. Short ubiquitin chains on CMG are sufficient to recruit the NEIL3 glycosylase, while longer chains enable p97-dependent CMG unloading and activation of the FA pathway. We confirmed that colibactin-induced ICL repair requires fork convergence (Supplementary Fig. 2.1i) and so tested whether it also requires CMG unloading. Addition of p97 inhibitor (NMS-873[38]) to pICL^{c742} and pICL^{clb} replication reactions caused a persistence of slow figure 8s and a reduction in unhooked open circular and supercoiled products (Fig. 2.1e), indicating that repair is activated by CMG unloading. Importantly, the absence of unhooking in p97-inhibited extracts implies that c742- and clb-ICLs are not processed by the NEIL3 or AA-ICL pathways, which do not require CMG unloading[14,15]. We additionally confirmed that depletion of NEIL3 from extracts has a negligible effect on the repair of pICL^{c742} (Supplementary Fig. 2.1j,k). We also examined the effect of p97 inhibition on maturation of nascent DNA strands by digesting repair

intermediates with restriction enzymes (AflIII and EcoRI) that enable resolution of nascent leading strands on a denaturing sequencing gel (Fig. 2.1f). During pICL^{c742} and pICL^{clb} replication, the nascent leading strands initially stall ~20 to 40 nucleotides (nt) upstream of the ICL due to the footprint of CMG. They are subsequently extended to 1 nt upstream of the expected dA cross-link positions before eventually being converted into full-length extension products. Consistent with a requirement for CMG unloading during pICL^{c742} and pICL^{clb} repair, p97 inhibitor attenuated extension of the nascent strands from the -20 to -40 stall positions to the -1 position and suppressed the formation of extension products. Note that the effect on extension product accumulation is more pronounced for pICL^{c742} than for pICL^{clb}, consistent with the presence of uncross-linked plasmid in the pICL^{clb} preparation. We conclude that colibactin-induced ICL repair is activated by fork convergence and CMG unloading, suggestive of FA pathway repair.

We noticed that replication of pICL^{c742} consistently produced two prominent intermediates that were not observed during replication of pICL^{AP}, pICL^{Pt}, or pICL^{clb} (e.g. Fig. 2.1d,e, red arrowheads). Our characterizations of these intermediates indicate that both species accumulate after CMG unloading and approach of the nascent leading strands up to the ICL (Fig. 2.1e and Supplementary Fig. 2.1h; see below). We suspect that these intermediates correspond to topological isomers of cross-linked fast figure 8 structures that have undergone CMG unloading but not fork reversal. It is possible that the structure of the c742-ICL might somehow antagonize fork reversal, though additional studies will be required to confirm this hypothesis.

2.3.2 FA pathway proteins are recruited to clb-ICLs during replication

To determine which repair networks are activated during colibactin-induced ICL repair, we performed unbiased quantitative mass spectrometry on chromatin recovered from pICL^{c742} replication reactions (Supplementary Fig. 2.2a,b)[39,40]. As expected, we observed that the CMG helicase, Pol δ, Pol ε, and other replisome components were enriched on pICL^{c742} at early replication time points (Fig. 2.2a). Replisome recruitment was inhibited by geminin, which blocks helicase loading (Fig. 2.2a and Supplementary Fig. 2.2c)[41]. Replisome components decreased in abundance over time, consistent with both replication termination and replisome disassembly following fork convergence at the ICL (Fig. 2.2a and Supplementary Fig. 2.2d). Replisome dissociation from pICL^{c742} coincided with recruitment of proteins that participate in the FA pathway, including ATR-ATRIP, FANCI-FANCD2, structure-specific nuclease complexes, TLS polymerases, and homologous recombination proteins (Fig. 2.2a). Importantly, FA pathway proteins were enriched on pICL^{c742} relative to both pCtrl and pICL^{c742} that was incubated in egg extract supplemented with geminin, indicating that these repair factors are recruited in a damage- and replication-dependent manner (Fig. 2.2b,c). Surprisingly, FANCM and the FA core complex were most abundant on pCtrl (at both 10 and 90 min after replication initiation) and on pICL^{c742} that was incubated in geminin-supplemented extract (Fig. 2.2a). These factors were similarly abundant during the first 30 minutes of pICL^{c742} replication. Despite the puzzling recruitment dynamics of the FA core complex, we nonetheless observed that replication of pICL^{c742} and pICL^{clb}, but not pCtrl, stimulated mono-ubiquitylation of FANCD2 in egg extracts (Fig. 2.2d) and that this ubiquitylation was temporally correlated with recruitment

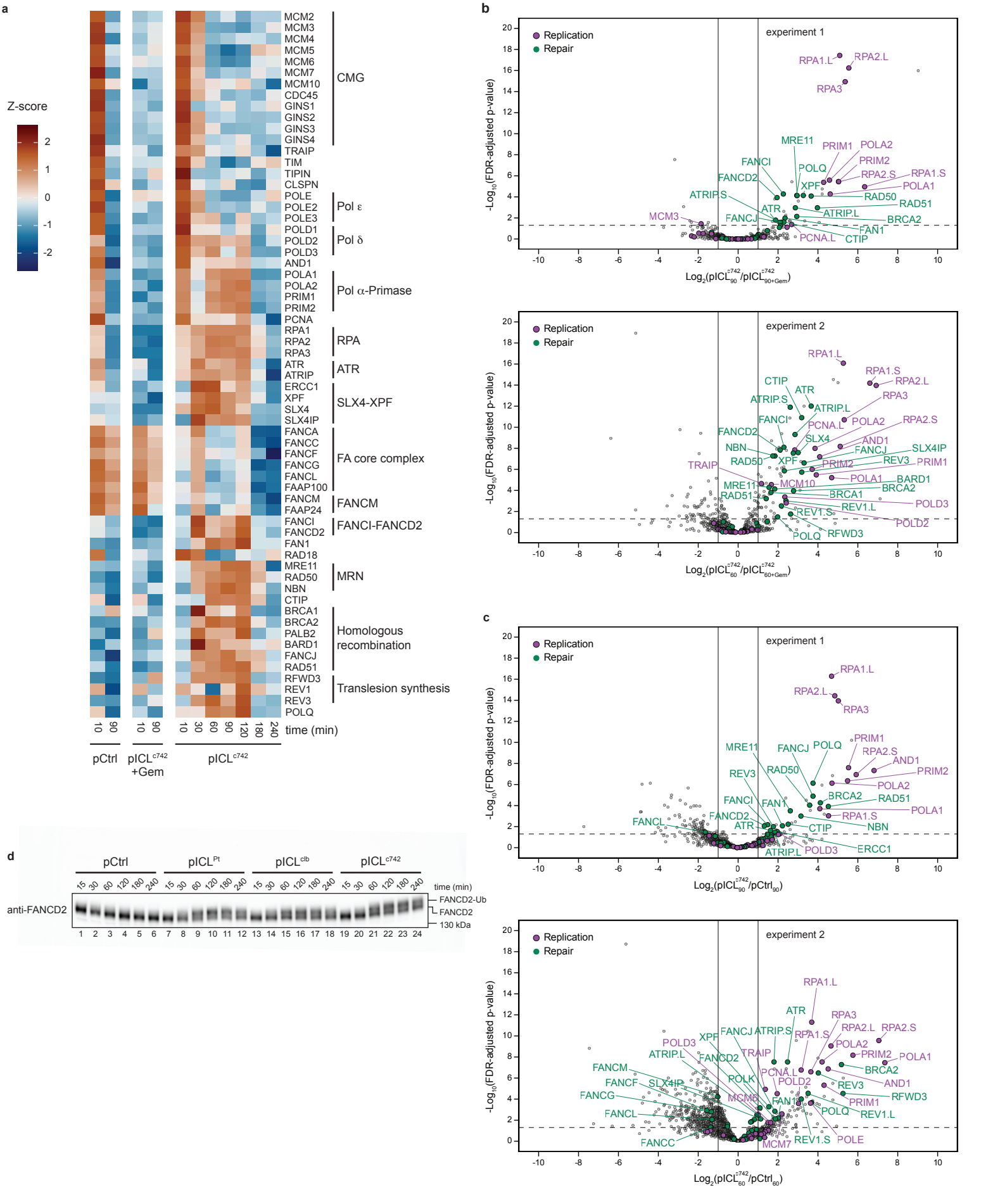

a
Z-score
2
1
0
-1
-2
MCM2
MCM3
MCM4
MCM5
MCM6
MCM7
MCM10
CDC45
GINS1
GINS2
GINS3
GINS4
TRAIP
TIM
TIPIN
CLSPN
POLE
POLE2
POLE3
POLD1
POLD2
POLD3
AND1
POLA1
POLA2
PRIM1
PRIM2
PCNA
RPA1
RPA2
RPA3
ATR
ATRIP
ERCC1
XPF
SLX4
SLX4IP
FANCA
FANCC
FANCF
FANCG
FANCL
FAAP100
FANCM
FAAP24
FANCI
FANCD2
FAN1
RAD18
MRE11
RAD50
NBN
CTIP
BRCA1
BRCA2
PALB2
BARD1
FANCJ
RAD51
RFWD3
REV1
REV3
POLQ
CMG
Pol ε
Pol δ
Pol α-Primase
RPA
ATR
SLX4-XPF
FA core complex
FANCM
FANCI-FANCD2
MRN
Homologous recombination
Translesion synthesis
10 90 10 90 10 30 60 90 120 180 240
time (min)
pCtrl
pICL c742 +Gem
pICL c742
b
Replication
Repair
experiment 1
-Log10(FDR-adjusted p-value)
Log2(pICL c742 90 / pICL c742 90+Gem)
RPA1.L
RPA2.L
RPA3
MRE11
PRIM1
POLA2
FANCI
POLQ
PRIM2
FANCD2
RPA2.S
RPA1.S
XPF
POLA1
ATRIP.S
RAD50
ATR
RAD51
MCM3
FANCJ
ATRIP.L
BRCA2
FAN1
PCNA.L
CTIP
experiment 2
Log2(pICL c742 60 / pICL c742 60+Gem)
RPA1.L
RPA1.S
RPA2.L
CTIP
ATR
ATRIP.S
RPA3
ATRIP.L
FANCI
PCNA.L
POLA2
RPA2.S
FANCD2
SLX4
AND1
NBN
FANCJ
SLX4IP
RAD50
XPF
PRIM2
REV3
PRIM1
TRAIP
POLA1
BARD1
MRE11
MCM10
BRCA2
RAD51
BRCA1
POLD3
REV1.L
REV1.S
POLD2
POLQ
RFWD3
c
Replication
Repair
experiment 1
-Log10(FDR-adjusted p-value)
Log2(pICL c742 90 / pCtrl 90)
RPA1.L
RPA2.L
RPA3
PRIM1
MRE11
FANCJ
POLQ
RPA2.S
AND1
REV3
RAD50
PRIM2
POLA2
BRCA2
FANCI
FAN1
RAD51
POLA1
FANCL
FANCD2
RPA1.S
ATR
CTIP
NBN
ATRIP.L
POLD3
ERCC1
experiment 2
Log2(pICL c742 60 / pCtrl 60)
RPA1.L
ATR
RPA3
FANCJ
ATRIP.S
RPA1.S
RPA2.L
RPA2.S
XPF
POLD3
POLA2
ATRIP.L
FANCD2
TRAIP
PRIM2
POLA1
PCNA.L
BRCA2
FANCM
POLK
POLD2
REV3
AND1
FANCF
SLX4IP
FANCG
MCM8
FAN1
RFWD3
PRIM1
FANCL
POLQ
REV1.L
REV1.S
POLE
FANCC
MCM7
d
pCtrl
pICL Pt
pICL clb
pICL c742
15 30 60 120 180 240
time (min)
anti-FANCD2
FANCD2-Ub
FANCD2
130 kDa
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

Fig. 2.2 | FA pathway mobilization during pICL^{c742} replication

a, pCtrl and pICL^{c742} were replicated in egg extract. Where indicated, extract was supplemented with Geminin (Gem) to inhibit replication. Plasmids were recovered from extract, and chromatin-associated proteins were quantified by mass spectrometry. The heat map shows the Z-scored median log_2LFQ (label-free quantification) intensity from three samples collected from the same replication reaction (i.e. technical replicates).

b, Volcano plots comparing the recruitment of proteins to pICL^{c742} in replicating extract (pICL) vs non-replicating extract (pICL$_{+Gem}$) at 90 minutes (experiment 1, top) or 60 minutes (experiment 2, bottom) following replication initiation. The estimated fold change (FC) in normalized protein intensity between the samples is plotted against false discovery rate- (FDR-)adjusted p-value (calculated with a two-sided moderated t-test and corrected for multiple hypothesis testing). Labeled proteins of interest meet significance criteria ($|Log_2FC| \geq 1$ and $p \leq 0.05$).

c, Volcano plots comparing the recruitment of proteins to pICL^{c742} vs pCtrl at 90 minutes (experiment 1, top) or 60 minutes (experiment 2, bottom) following replication initiation. Estimated fold changes and FDR-adjusted p-values were calculated as in **b**. Labeled proteins of interest meet significance criteria ($|Log_2FC| \geq 1$ and $p \leq 0.05$).

d, The indicated plasmids were replicated in egg extract, and extract proteins were resolved by SDS-PAGE and blotted for FANCD2. In the presence of an ICL, FANCD2 becomes monoubiquitylated, resulting in a slower migrating band.

of FANCI-FANCD2 to chromatin (Fig. 2.2a). Taken together, these data indicate that the FA pathway is mobilized during replication-coupled repair of colibactin-induced ICLs.

2.3.3 Clb-ICLs are unhooked by the FA pathway

ICL repair by the FA pathway involves nucleolytic incisions that effectively convert the ICL into a DSB. Replication of pICL^{c742} resulted in the transient accumulation of DNA fragments consistent with DSB formation during nucleolytic repair (Supplementary Fig.

2.3a). This accumulation of putative pICL^{c742} incision products closely mirrored the accumulation of known FA pathway pICL^{Pt} incision products[36,37]. By contrast, incision products did not appreciably accumulate during replication of pICL^{AP}, which is predominantly unhooked without incisions by NEIL3[15]. Over time, the incision products disappeared, presumably due to DSB repair by homologous recombination factors, which our data indicate are recruited to plasmids containing a colibactin-induced ICL (Fig. 2.2).

To explicitly test whether replication-coupled repair of colibactin-induced ICLs requires the FA pathway, we replicated pICL^{c742} and pICL^{clb} in mock- or FANCD2-immunodepleted extracts supplemented with recombinant (r)FANCI-FANCD2, as indicated (Fig. 2.3a and Supplementary Fig. 2.3b,c). Native agarose gel analysis of replication intermediates revealed that FANCD2 depletion resulted in an accumulation of reversed fork intermediates and concomitant reductions in unhooked open circular and supercoiled plasmids (Fig. 2.3b and Supplementary Fig. 2.3d), indicating a defect in ICL repair. Note that the effects of FANCD2 depletion on the accumulation of open circular and supercoiled plasmids were more subtle during replication of pICL^{clb} due to the presence of uncross-linked plasmid. Importantly, the accumulation of unhooked open circular and supercoiled plasmids was restored by addition of rFANCI-FANCD2^{WT} complex, but not by addition of rFANCI-FANCD2^{K562R}, which cannot undergo FANCD2 mono-ubiquitylation[22]. These results indicate that repair of colibactin-induced ICLs depends on ubiquitylated FANCI-FANCD2. To confirm that FANCI-FANCD2 is required to unhook colibactin-induced ICLs, pICL^{c742} was replicated in mock- or FANDC2-depleted extract, and recovered replication intermediates and products were then linearized, resolved on a denaturing agarose gel, and detected by Southern blotting. FANCD2 depletion caused a persistence of

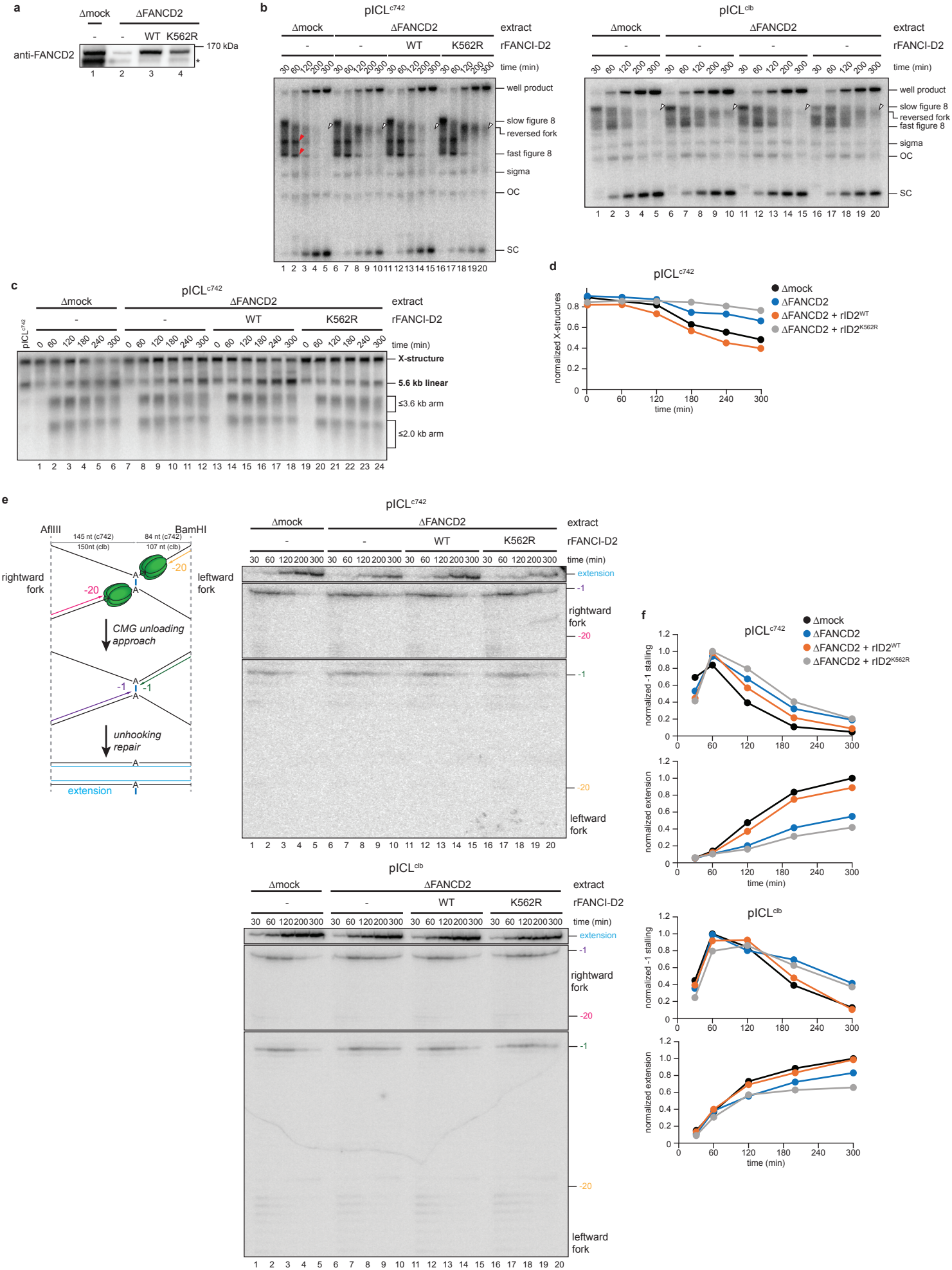

a
Δmock
ΔFANCD2
- - WT K562R
anti-FANCD2
170 kDa
1 2 3 4
b
pICL^c742
pICL^clb
extract
rFANCI-D2
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
c
pICL^c742
X-structure
5.6 kb linear
≤3.6 kb arm
≤2.0 kb arm
d
pICL^c742
normalized X-structures
time (min)
Δmock
ΔFANCD2
ΔFANCD2 + rID2^WT
ΔFANCD2 + rID2^K562R
e
AflIII
BamHI
145 nt (c742)
84 nt (c742)
150nt (clb)
107 nt (clb)
rightward fork
leftward fork
CMG unloading approach
unhooking repair
extension
pICL^c742
pICL^clb
f
normalized -1 stalling
normalized extension
time (min)

Fig. 2.3 | FA pathway-dependent unhooking of colibactin-induced ICLs

a, FANCD2 immunodepletion. The egg extracts used in the replication reactions shown in **b** (left) and **e** (top) were blotted for FANCD2. Asterisk indicates non-specific band.

b, pICL^{c742} (left) and pICL^{clb} (right) were replicated with [α-^{32}P]dCTP in mock- or FANCD2-depleted egg extracts supplemented with rFANCI-FANCD2^{WT} or rFANCI-FANCD2^{K562R}, as indicated. Replication intermediates were analyzed as in Fig. 2.1d. SC, supercoiled; OC, open circular. Red arrowheads indicate pICL^{c742}-specific repair intermediates; white arrowheads indicate reversed fork intermediates, which accumulate upon depletion of FANCD2.

c, pICL^{c742} was replicated in mock- or FANCD2-depleted egg extracts supplemented with rFANCI-FANCD2^{WT} or rFANCI-FANCD2^{K562R}, as indicated. Replication intermediates were analyzed by Southern blotting as in Fig. 2.1c.

d, Quantification of the Southern blot shown in **c**. The persistence of pICL^{c742} X-structures was quantified as the intensity of X-structures relative to the combined intensities of X-structures (X) and 5.6 kb linear species (L; X/X+L).

e, Left, schematic of nascent leading strand stall and extension products generated during ICL repair by the FA pathway, as in Fig. 2.1f. Right, nascent DNA strands from the pICL^{c742} (top) and pICL^{clb} (bottom) replication reactions shown in **b** were digested with AflIII and BamHI and analyzed as in Fig. 2.1f.

f, Quantification of the -1 stall and strand extension products shown in **e**. The intensities of rightward fork -1 stall products (top) or extension products (bottom) were normalized to maximum -1 stall or extension product intensity, respectively, across all samples.

cross-linked X-structures and a reduction in unhooked linear plasmids (Fig. 2.3c,d and Supplementary Fig. 2.3e,f). The conversion of X-structures into linear plasmids was restored by addition of rFANCI-FANCD2^{WT}, but not by addition of FANCI-FANCD2^{K562R}. These data indicate that colibactin-induced ICLs are resolved through nucleolytic incisions mediated by FANCI-FANCD2.

Inhibition of ICL unhooking by the FA pathway is also expected to block extension of nascent DNA strands past the ICL. Analysis of the nascent DNA leading strands revealed that FANCD2 depletion caused a persistence of -1 stall products and a reduction in full-length extension products that was rescued by rFANCI-FANCD2^{WT}, but not by rFANCI-FANCD2^{K562R} (Fig. 2.3e,f). As has been observed during cisplatin-ICL repair (Supplementary Fig. 2.3g-j)[37], FANCI-FANCD2 is therefore required for extension of nascent leading strands past the colibactin-induced ICL. Taken together, our data indicate that, following CMG unloading and fork reversal, FANCI-FANCD2 promotes nucleolytic incisions that unhook the cross-linked parental DNA strands, thereby enabling TLS past the resulting ICL remnant and completion of DNA replication.

2.3.4 TLS during clb-ICL repair depends on REV1, Pol κ, and Pol η

Following ICL unhooking, TLS extends the nascent DNA strands past damaged nucleotides in the template DNA strands. For ICLs formed by nitrogen mustard and cisplatin, TLS is mediated by an insertion polymerase, which inserts a nucleotide opposite the damaged base, and the REV1-Pol ζ complex, which extends the DNA strand[23]. In contrast, REV1 is required for both insertion and extension during repair of ICLs formed by acetaldehyde, psoralen, or AP sites[14,15]. REV1 immunodepletion largely abolished the formation of supercoiled plasmids during pICL^{c742} replication (Fig. 2.4a,b). This was accompanied by an accumulation of open circular plasmids, indicating a requirement for REV1 during TLS past the unhooked c742-ICL remnant. REV1 depletion also caused a pronounced accumulation of open circular plasmids during pICL^{clb} replication (Fig. 2.4b), indicating that bypass of either an unhooked clb-ICL remnant or a clb-A monoadduct also depends on REV1 (see below). Supercoiled plasmid still accumulated during pICL^{clb}

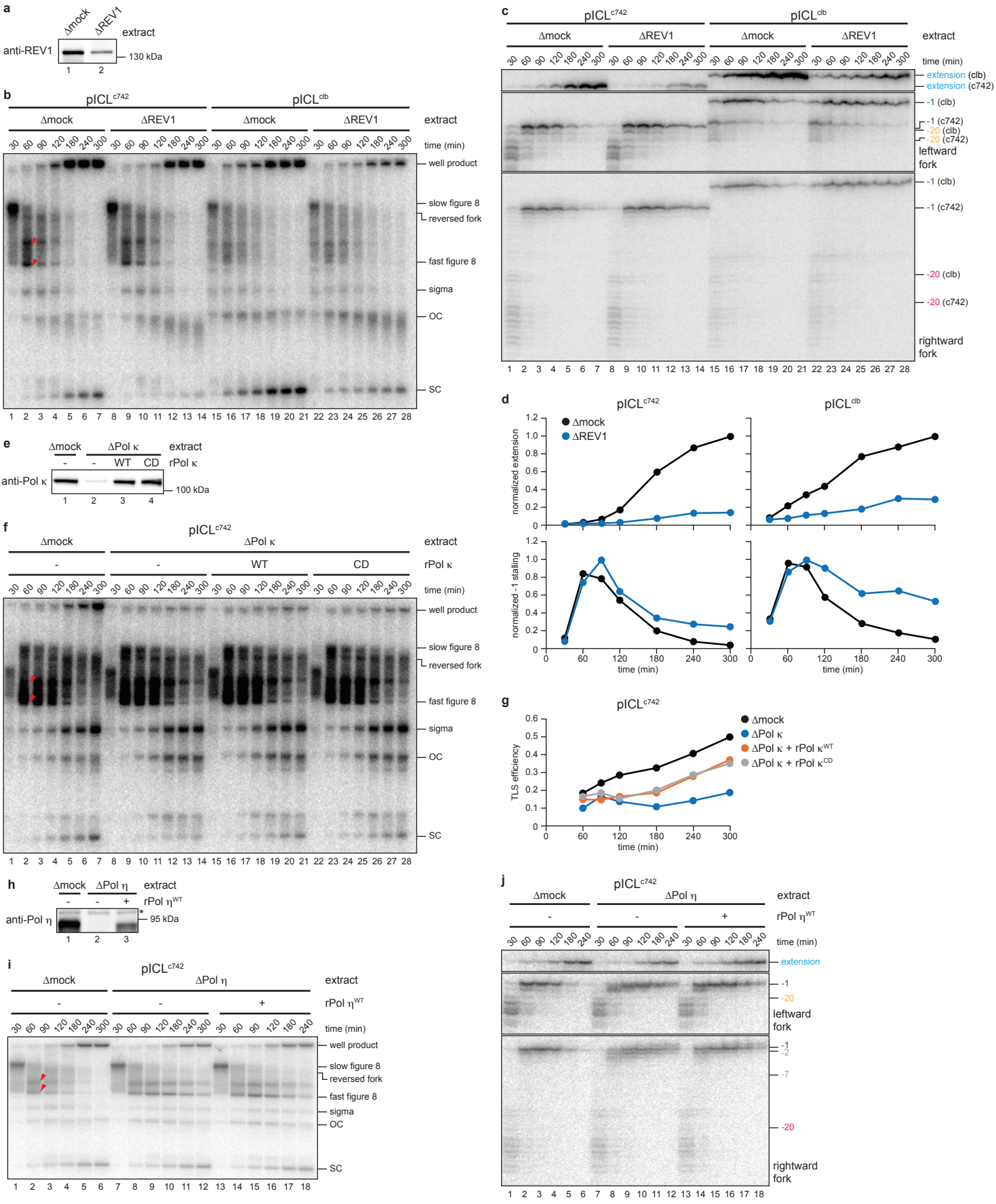

a
Δmock
ΔREV1
extract
anti-REV1
130 kDa
1 2
b
pICL^c742
pICL^clb
Δmock
ΔREV1
extract
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
c
pICL^c742
pICL^clb
extension (clb)
extension (c742)
-1 (clb)
-1 (c742)
-20 (clb)
-20 (c742)
leftward fork
rightward fork
d
pICL^c742
pICL^clb
Δmock
ΔREV1
normalized extension
normalized -1 stalling
time (min)
e
Δmock
ΔPol κ
extract
rPol κ
WT
CD
anti-Pol κ
100 kDa
f
pICL^c742
Δmock
ΔPol κ
rPol κ
g
pICL^c742
Δmock
ΔPol κ
ΔPol κ + rPol κ^WT
ΔPol κ + rPol κ^CD
TLS efficiency
time (min)
h
Δmock
ΔPol η
extract
rPol η^WT
anti-Pol η
95 kDa
i
pICL^c742
Δmock
ΔPol η
extract
rPol η^WT
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
j
pICL^c742
Δmock
ΔPol η
extract
rPol η^WT
time (min)
extension
-1
-20
leftward fork
-1
-2
-7
-20
rightward fork

Fig. 2.4 | Translesion synthesis during colibactin-induced ICL repair

a, REV1 immunodepletion. The extracts used in the replication reactions shown in **b** and **c** were blotted for REV1.

b, pICL^{c742} and pICL^{clb} were replicated with [α-^{32}P]dCTP in mock- or REV1-depleted extracts, and replication intermediates were analyzed as in Fig. 2.1d. SC, supercoiled; OC, open circular. Red arrowheads indicate pICL^{c742}-specific repair intermediates.

c, Nascent DNA strands from the pICL^{c742} and pICL^{clb} replication reactions shown in **b** were digested with AflIII and EcoRI and analyzed as in Fig. 2.1f.

d, Rightward fork -1 stall products and extension products for the replication reactions shown in **c** were quantified as in Fig. 2.3f.

e, Pol κ immunodepletion. The extracts used in the replication reactions shown in **f** were blotted for Pol κ.

f, pICL^{c742} was replicated with [α-^{32}P]dCTP in mock- or Pol κ-depleted extracts supplemented with rPol κ, as indicated, and replication intermediates were analyzed as in Fig. 2.1d. SC, supercoiled; OC, open circular. Red arrowheads indicate top and bottom intermediates formed during replication of pICL^{c742}.

g, TLS efficiencies for the replication reactions shown in **f** were quantified as the intensity of supercoiled plasmid relative to combined intensities of open circular and supercoiled plasmids (SC/OC+SC).

h, Pol η immunodepletion. The extracts used in the replication reactions shown in **i** and **j** were blotted for Pol η. Asterisk indicates non-specific band.

i, pICL^{c742} was replicated with [α-^{32}P]dCTP in mock- or Pol η-depleted extracts supplemented with rPol η, as indicated. Replication intermediates were analyzed as in Fig. 2.1d. SC, supercoiled; OC, open circular. Red arrowheads indicate pICL^{c742}-specific repair intermediates.

j, Nascent DNA strands from the pICL^{c742} replication reactions shown in **i** were digested with AflIII and EcoRI and analyzed as in Fig. 2.1f.

replication in REV1-depleted extract, likely due to repair of template strand AP sites produced upon clb-ICL depurination (Supplementary Fig. 2.1f). For both pICL^{c742} and pICL^{clb}, REV1 depletion caused a strong reduction in nascent strand extension products and a persistence of -1 stall products (Fig. 2.4c,d and Supplementary Fig. 2.4a), distinct

from the persistence of insertion (0 stall position) products observed during cisplatin-ICL repair in REV1-depleted extracts (Supplementary Fig. 2.4b-d)[23]. Our data therefore indicate that REV1 is required for both insertion and extension during TLS past a colibactin-induced ICL remnant. In some experiments, small amounts of insertion products were detected during repair of the c742-ICL that were suppressed by REV1 depletion (e.g. Supplementary Fig. 2.4d). These insertion products were not detected during bypass of the clb-ICL, suggesting that differences between the structures of the c742- and clb-ICLs may influence the efficiency of REV1-dependent insertion opposite an alkylated dA.

REV1-Pol ζ has been shown to mediate TLS in concert with Pol κ, which functions non-catalytically to stabilize the REV1-Pol ζ polymerase complex[42]. We tested whether Pol κ also promotes TLS during repair of pICL^{c742}. Like REV1 immunodepletion, Pol κ immunodepletion strongly inhibited formation of supercoiled plasmids and caused a corresponding accumulation of open circular plasmid, consistent with defective TLS (Fig. 2.4e-g). Addition of either wild-type (rPol κ^{WT}) or D198A/E199A mutated, catalytically dead (rPol κ^{CD}) rPol κ partially restored the formation of supercoiled plasmids, indicating that Pol κ acts non-catalytically during TLS past a c742-ICL remnant (Fig. 2.4e-g and Supplementary Fig. 2.4e,f). The inability of rPol κ to fully rescue the TLS defect caused by Pol κ immunodepletion might be attributed to co-depletion of other TLS factors. For example, in previous work, the Pol κ antibody used in our experiments co-depleted the FA core complex, which has been implicated in TLS regulation[42,43]. Overall, our data indicate that TLS during colibactin-induced ICL repair is mediated by a Pol κ-REV1-Pol ζ complex.

We also tested whether TLS during colibactin-induced ICL repair depends on Pol η, which has been reported to function as an insertion polymerase during TLS past ICLs and other

DNA lesions[44]. Pol η immunodepletion did not appreciably affect the formation of supercoiled plasmids during pICL^{c742} replication in egg extracts (Fig. 2.4h,i), suggesting that Pol η is dispensable for TLS during colibactin-induced ICL repair. However, nascent strand analysis revealed that Pol η depletion did cause a transient accumulation of stall products from the -2 to the -7 position, which was suppressed by addition of wild-type rPol η (Fig. 2.4j). This pattern of stall product accumulation resembled that observed during pICL^{Pt} replication (Supplementary Fig. 2.4b-d). We conclude that Pol η normally extends nascent leading strands to within 1 nt of a colibactin-induced ICL. When Pol η is absent, however, another DNA polymerase – such as Pol δ – can perform this function, albeit less efficiently.

2.3.5 TLS past a colibactin-induced monoadduct depends on REV1 and Pol κ

Previous work has implicated the polymerase activity of Pol κ in TLS past minor groove adducts formed upon N3 dA alkylation[42,45]. We therefore wondered whether the catalytic activity of Pol κ is required for TLS past a colibactin-induced N3 dA adduct outside the context of ICL repair. During preparation of pICL^{clb}, depurination of the clb-ICL resulted in the formation of site-specific clb-A monoadducts and AP sites, the latter of which are rapidly removed in egg extract (Supplementary Fig. 2.1d-f)[34]. We exploited this decomposition of the clb-ICL to examine the role of Pol κ during TLS past a colibactin-induced monoadduct, independent of ICL unhooking. pICL^{clb} was replicated in mock- or Pol κ-depleted egg extract that was supplemented with p97 inhibitor to block CMG unloading and suppress approach of the nascent leading strands up to the ICL (Fig. 2.5a and Supplementary Fig. 2.5a). Under these conditions, TLS occurs only upon encounters with clb-A monoadducts produced by depurination. Pol κ depletion inhibited the formation

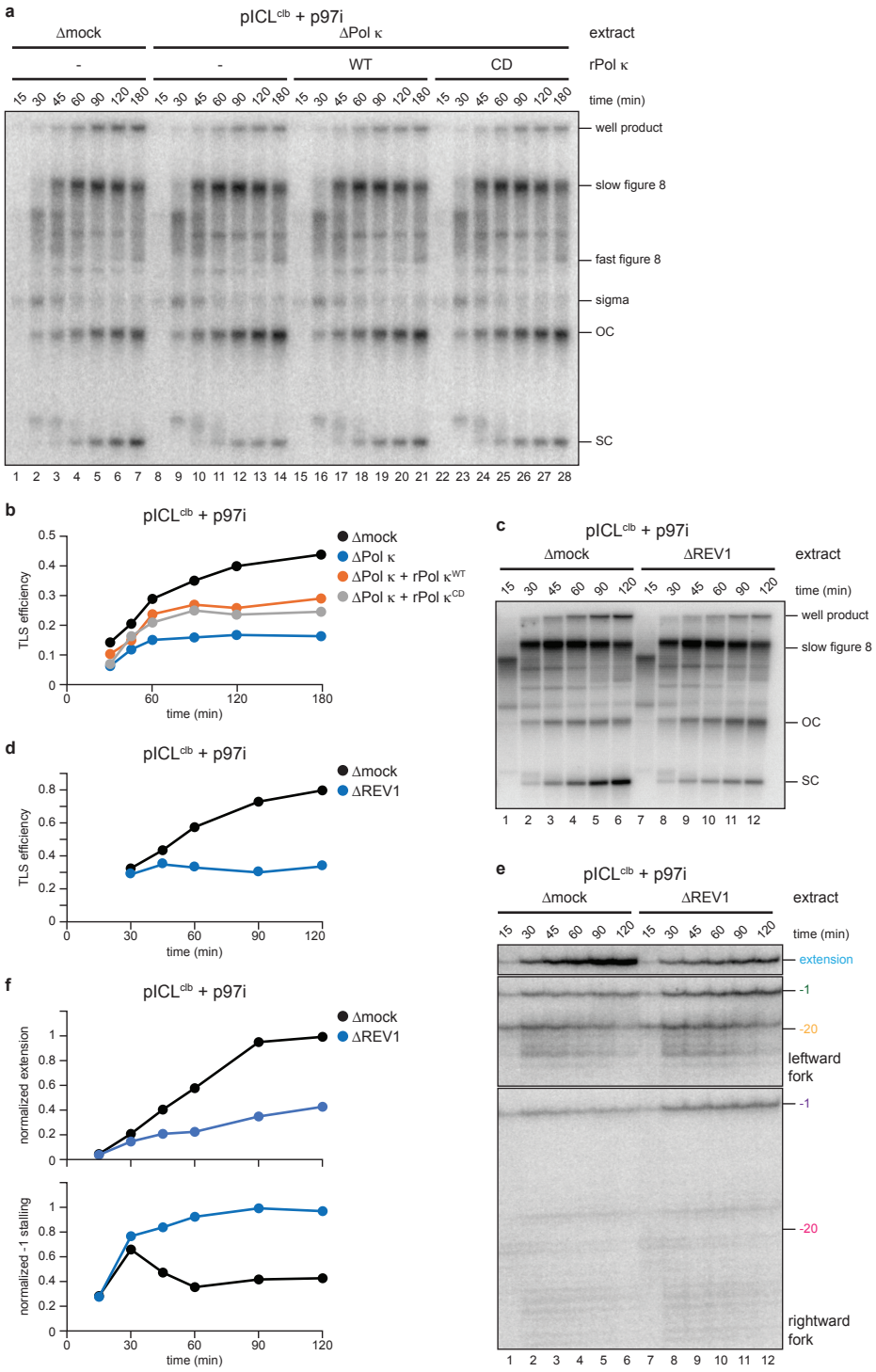


Fig. 2.5 | Translesion synthesis during bypass of a colibactin-induced monoadduct

a, pICL^{clb} was replicated with [α-^{32}P]dCTP in mock- or Pol κ-depleted extracts supplemented with p97 inhibitor and rPol κ, as indicated, and replication intermediates were analyzed as in Fig. 2.1d. SC, supercoiled; OC, open circular.

b, TLS efficiencies for the replication reactions shown in **a** were quantified as in Fig. 2.4g.

c, pICL^{clb} was replicated with [α-^{32}P]dCTP in mock- or REV1-depleted extracts supplemented with p97 inhibitor, and replication intermediates were analyzed as in Fig. 2.1d. SC, supercoiled; OC, open circular.

d, TLS efficiencies for the replication reactions shown in **c** were quantified as in Fig. 2.4g.

e, Nascent DNA strands from the pICL^{clb} replication reactions shown in **c** were digested with AflIII and EcoRI and analyzed as in Fig. 2.1f.

f, Rightward fork -1 stall products and extension products for the replication reactions shown in **e** were quantified as in Fig. 2.3f.

of fully replicated, supercoiled plasmids and caused a corresponding accumulation of open circular plasmids, indicating that TLS past the clb-A adduct requires Pol κ (Fig. 2.5a,b and Supplementary Fig. 2.5b). However, the formation of supercoiled plasmids was partially rescued by addition of either rPol κ^{WT} or rPol κ^{CD}, indicating that the catalytic activity of Pol κ is dispensable for TLS past the clb-A adduct.

These results suggest that Pol κ acts non-catalytically to promote TLS in complex with REV1-Pol ζ. During pICL^{clb} replication in extract supplemented with p97 inhibitor, REV1 depletion inhibited the formation of supercoiled plasmids (Fig. 2.5c,d and Supplementary Fig. 2.5c,d) and caused persistent stalling of nascent leading strands at the -1 position (Fig. 2.5e,f and Supplementary Fig. 2.5e), implying that REV1-Pol ζ promotes TLS past a clb-A monoadduct. Our data therefore indicate that unhooked c742-ICL remnants and clb-A monoadducts are bypassed using a common mechanism that depends on the Pol κ-REV1-Pol ζ complex.

2.3.6 Bypass of colibactin-induced DNA damage is mutagenic

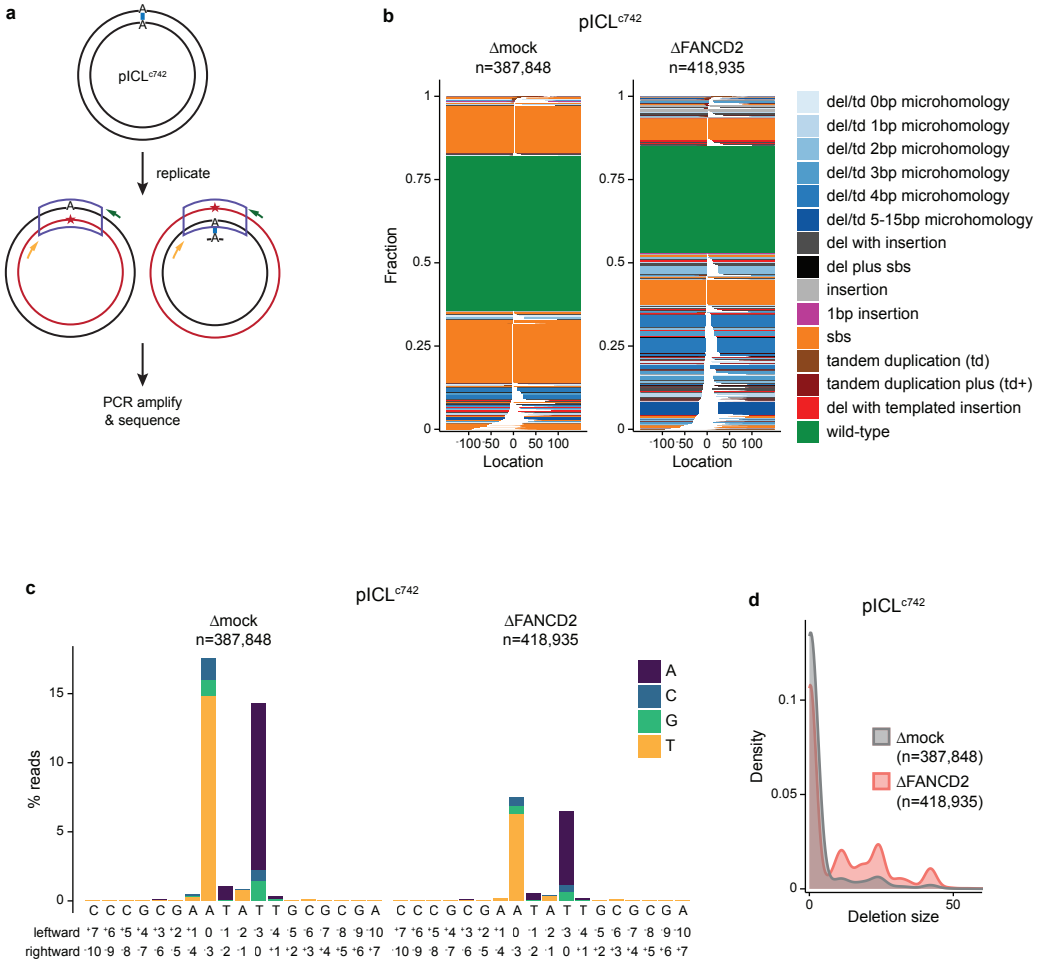


Fig. 2.6 | Mutagenic repair of colibactin-induced ICLs by the FA pathway

a, Schematic depicting analysis of pICL^{c742} repair products by next-generation sequencing. Blue and green arrows indicate PCR primers; red asterisk indicates potential point mutation.

b, pICL^{c742} was replicated in mock- or FANCD2-depleted extract, and repair products were PCR amplified and sequenced. Raw paired-end sequencing reads were merged using SIQ, and data were visualized with SIQPlotteR. Color-coded repair outcomes are depicted as tornado plots. White spaces depict deletions. The positions of variations are indicated relative to the AATATT cross-link site ("0"). n indicates the number of mapped sequencing reads obtained for each condition.

c, The percentage of reads containing single base substitutions at the indicated positions relative to the ICL is plotted for the data sets described in **b**. The reference sequence is shown with numbers indicating the positions for the rightward and leftward nascent leading strands relative to the site of cross-linking ("0"). n indicates the number of mapped sequencing reads obtained for each condition.

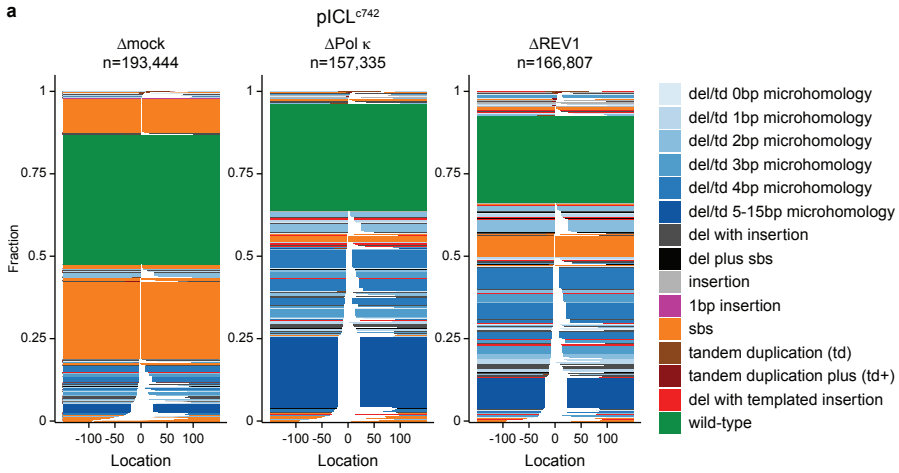
d, The fraction of deletions spanning a given length is plotted for the data sets described in **b**. n indicates the number of mapped sequencing reads obtained for each condition.

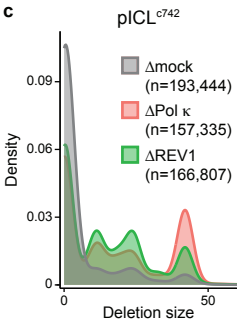
Exposure to *pks*$^{+}$ bacteria produces T>N SBSs and single nucleotide T deletions at colibactin alkylation hotspots[4-6,10], but the repair processes responsible for these mutational signatures have not been established. To determine whether colibactin-induced ICL repair by the FA pathway is mutagenic, we replicated pICL^{c742} in mock- or FANCD2-depleted egg extract and performed next-generation sequencing on recovered repair products (Fig. 2.6a and Supplementary Fig. 2.6a-c). Following replication in mock-depleted extract, we observed that 46.6% of the plasmid sequences were free of mutations, insertions, and deletions, while 34.5% of plasmid sequences contained SBSs within the AATATT colibactin hotspot (Fig. 2.6b,c, left). The vast majority (~92%) of these point mutations mapped to the expected alkylation sites (Fig. 2.6c). These data indicate that while bypass of colibactin-induced DNA damage can be error-free, it is frequently mutagenic. In cells and organoids, colibactin exposure is associated primarily with T>C SBSs, and less frequently with T>A and T>G SBSs. Interestingly, c742-ICL repair produced mostly T>A SBSs, with fewer T>G and T>C SBSs. This suggests that while the FA pathway may account for colibactin-induced T>A mutations, another mutagenic process may generate T>C mutations associated with colibactin exposure. Depletion of FANCD2 resulted in a reduction in the proportion of plasmid sequences corresponding to error-free repair products or products with hotspot SBSs (32.3% error-free, 15.3% hotspot SBS), indicating that the FA pathway enables mutagenic bypass of colibactin-induced ICLs (Fig. 2.6b,c; right). However, FANCD2-depletion did not completely abolish recovery of the error-free and SBS repair products. Since a fraction of our pICL^{c742} plasmid has undergone ICL

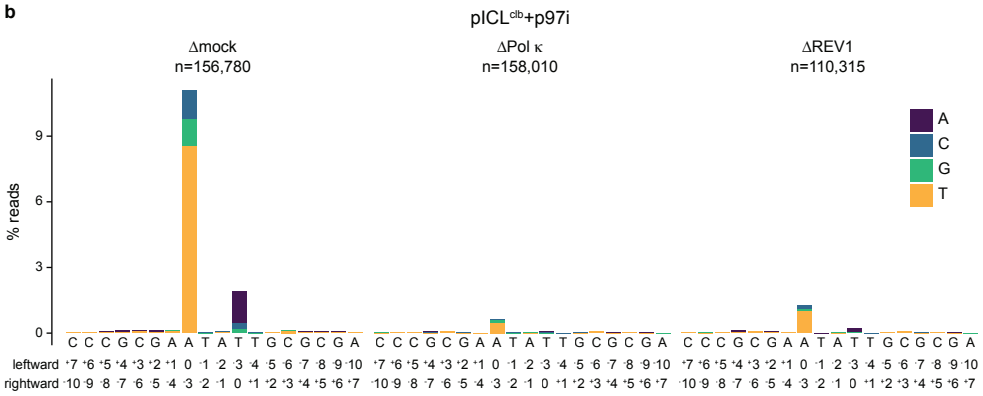
decomposition into a monoadduct and AP site (~25%, Supplementary Fig. 2.1c), we suspect that the remaining error-free and SBS repair products generated in FANCD2-depleted extract were derived from replication of initially uncross-linked plasmids.

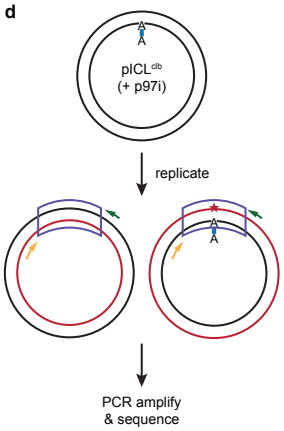
A substantial proportion of plasmid sequences produced during pICL^{c742} replication in mock-depleted extract contained deletion mutations (14%) (Fig. 2.6b, left). Essentially no insertion mutations were observed. The majority of deletion mutations (90.1%) at least partially overlapped the 4-bp region spanned by the cross-link. FANCD2-depletion resulted in a dramatic increase in the proportion of plasmid sequences containing deletion mutations (45.7%; Fig. 2.6b, right). As in mock-depleted extract, most of these deletions (91.7%) encompassed some part of the cross-linked region. Deletions produced during replication in FANCD2-depleted extract tended to be longer than those produced in mock-depleted extract (Fig. 2.6d). Similar trends were observed in two additional independent experiments (Supplementary Fig. 2.6d-f). We conclude that the FA pathway promotes colibactin-induced ICL repair that is generally error-free or introduces a point mutation at the site of alkylation. In the absence of the FA pathway, stalled replication forks may undergo aberrant fork cleavage and subsequent end-joining, resulting in the formation of deletions.

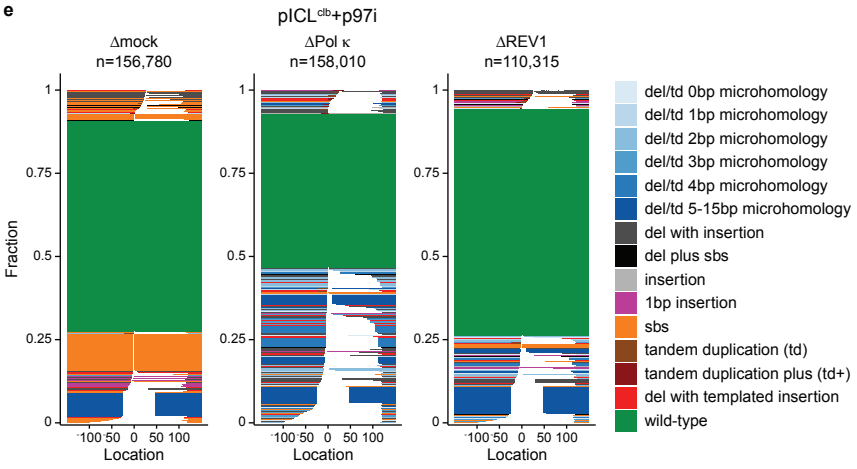
We next tested whether REV1 or Pol κ influence the spectrum of mutations observed following pICL^{c742} replication in egg extract. pICL^{c742} was replicated in mock-, REV1-, or Pol κ-depleted egg extract, and the repair products were sequenced (Supplementary Fig. 2.7a-f). Replication in mock-depleted extract produced 39.3% error-free repair products and 38.9% SBS repair products (Fig. 2.7a, left). By contrast, REV1-depletion reduced the proportion of plasmid sequences that were free of mutations or contained SBSs (26.4%

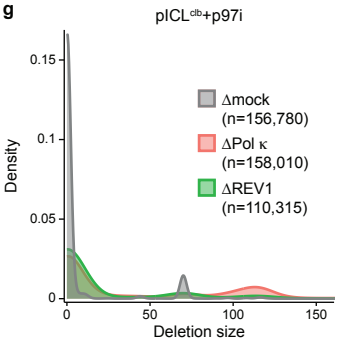

a
pICL c742
Δmock
n=193,444
ΔPol κ
n=157,335
ΔREV1
n=166,807
Fraction
Location
del/td 0bp microhomology
del/td 1bp microhomology
del/td 2bp microhomology
del/td 3bp microhomology
del/td 4bp microhomology
del/td 5-15bp microhomology
del with insertion
del plus sbs
insertion
1bp insertion
sbs
tandem duplication (td)
tandem duplication plus (td+)
del with templated insertion
wild-type


c
pICL c742
Δmock
(n=193,444)
ΔPol κ
(n=157,335)
ΔREV1
(n=166,807)
Density
Deletion size


b
pICL clb+p97i
Δmock
n=156,780
ΔPol κ
n=158,010
ΔREV1
n=110,315
% reads
A
C
G
T
leftward
rightward


d
pICL clb
(+ p97i)
replicate
PCR amplify
& sequence


e
pICL clb+p97i
Δmock
n=156,780
ΔPol κ
n=158,010
ΔREV1
n=110,315
Fraction
Location
del/td 0bp microhomology
del/td 1bp microhomology
del/td 2bp microhomology
del/td 3bp microhomology
del/td 4bp microhomology
del/td 5-15bp microhomology
del with insertion
del plus sbs
insertion
1bp insertion
sbs
tandem duplication (td)
tandem duplication plus (td+)
del with templated insertion
wild-type


g
pICL clb+p97i
Δmock
(n=156,780)
ΔPol κ
(n=158,010)
ΔREV1
(n=110,315)
Density
Deletion size

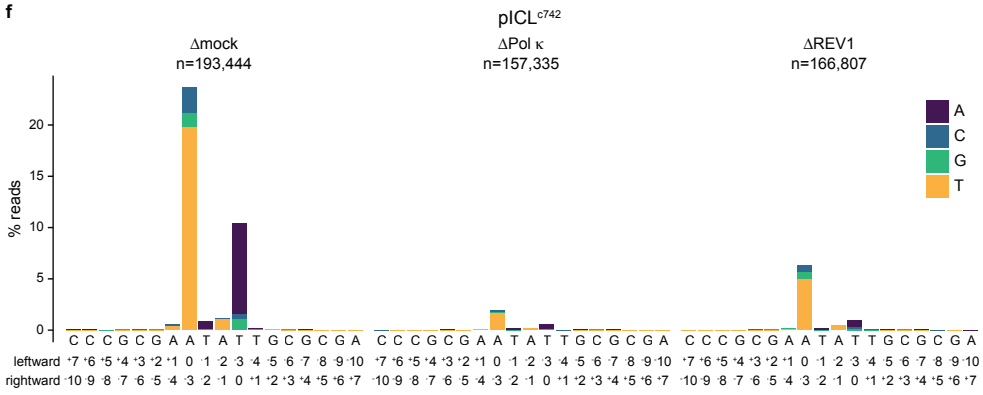

f
pICL c742
Δmock
n=193,444
ΔPol κ
n=157,335
ΔREV1
n=166,807
% reads
A
C
G
T
leftward
rightward

Fig. 2.7 | REV1 and Pol κ promote mutagenic bypass of colibactin-induced DNA damage

a, pICL^{c742} was replicated in mock-, Pol κ-, or REV1-depleted extract and repair products were PCR amplified, sequenced, and visualized as in Fig 6a,b. Color-coded repair outcomes are depicted as tornado plots, with white spaces indicating deletions. n indicates the number of mapped sequencing reads obtained for each condition.

b, The percentage of reads containing single base substitutions at the indicated positions relative to the ICL is plotted for the data sets described in **a**, as in Fig. 2.6c. n indicates the number of mapped sequencing reads obtained for each condition.

c, The fraction of deletions spanning a given length is plotted for the data sets described in **a**. n indicates the number of mapped sequencing reads obtained for each condition.

d, Schematic depicting analysis of pICL^{clb} repair products by next-generation sequencing. p97 inhibitor is added to block repair of intact clb-ICLs by the FA pathway. Blue and green arrows indicate PCR primers; red asterisk indicates potential point mutation.

e, pICL^{clb} was replicated in mock-, Pol κ-, or REV1-depleted extract supplemented with p97 inhibitor, and repair products were PCR amplified, sequenced, and visualized as in Fig 6a,b. Color-coded repair outcomes are depicted as tornado plots, with white spaces indicating deletions. n indicates the number of mapped sequencing reads obtained for each condition.

f, The percentage of reads containing single base substitutions at the indicated positions relative to the ICL is plotted for the data sets described in **e**, as in Fig. 2.6c. n indicates the number of mapped sequencing reads obtained for each condition.

g, The fraction of deletions spanning a given length is plotted for the data sets described in **e**. n indicates the number of mapped sequencing reads obtained for each condition.

error-free and 9.7% SBSs; Fig. 2.7a, right), as did Pol κ-depletion (32.4% error-free and 4.6% SBSs; Fig. 2.7a, center). Taken together, these results indicate that REV1 and Pol κ act in concert during mutagenic bypass of an unhooked colibactin-induced ICL. Notably, although the frequency of SBSs decreased upon REV1- or Pol κ-depletion, the relative distribution of T>A, T>G, and T>C mutations was similar across mock-, REV1-, and Pol

κ-depleted extracts (Fig. 2.7b), suggesting that residual REV1 and Pol κ mediate bypass of the ICL remnant in the depleted extracts. Strikingly, REV1- and Pol κ-depletion increased both the frequency and size of deletion mutations present in the recovered replication products: the frequency of reads containing deletions increased from 20.8% in mock-depleted extract to 61.5% in both REV1- and Pol κ-depleted extracts (Fig. 2.7a,c). These results suggest that an alternative deletion-prone mechanism can bypass colibactin-induced ICLs in the absence of REV1- and Pol κ-dependent TLS.

Since most fully replicated pICL^{clb} plasmids were generated upon replication of contaminating uncross-linked plasmids (Supplementary Fig. 2.1c-e), we were unable to distinguish the mutation spectrum produced during clb-ICL repair. Instead, we examined the mutagenicity of a colibactin-induced monoadduct. We replicated pICL^{clb} in mock-, REV1-, or Pol κ-depleted egg extract supplemented with p97 inhibitor (to once again inactivate ICL repair by the FA pathway) and sequenced the replication products (Fig. 2.7d and Supplementary Fig. 2.7g-l). Replication past clb-A monoadducts in mock-depleted extract produced a mutation spectrum mirroring that produced during c742-ICL repair. Error-free bypass was observed in 64.4% of the plasmid sequences, while SBSs were present in 16.8% of the plasmids sequences (Fig. 2.7e). Again, we observed predominantly T>A SBSs at the colibactin hotspot with smaller frequencies of T>G and T>C mutations (Fig. 2.7f). Depletion of either REV1 or Pol κ dramatically reduced the frequency of SBSs (5.8% and 3.0% for REV1- and Pol κ-depletion, respectively; Fig. 2.7e,f), indicating that REV1- and Pol κ-dependent TLS is responsible for the spectrum of point mutations observed upon bypass of a colibactin-induced monoadduct. REV1 and Pol κ depletions also resulted in a higher frequency of deletion mutations (14.5% in mock-depleted extract,

compared with 23.6% in REV1-depleted extract and 46.7% in Pol κ-depleted; Fig. 2.7e). This increase in deletion frequency was accompanied by a shift toward longer deletions (Fig. 2.7g). These trends were consistent across independent experimental replicates (Supplementary Fig. 2.7d-f,j-l). Thus, although REV1- and Pol κ-dependent TLS past a clb-A adduct introduces frequent point mutations, it appears to suppress an alternative bypass mechanism that is prone to larger deletions. Combined, our sequencing data indicate that REV1- and Pol κ-dependent TLS promotes mutagenic bypass of both unhooked colibactin-induced ICL remnants and colibactin-induced monoadducts.

2.4 Discussion

Despite intense interest in colibactin as a driver of colorectal cancer, little is known about the pathways that repair colibactin-induced DNA damage. Our data indicate that colibactin-induced ICL repair is activated upon convergence of replication forks at the lesion and unloading of the replisome (Fig. 2.8a). The FA pathway then mediates nucleolytic incisions that release the cross-link and generate a DSB, which is likely repaired by homologous recombination, and an ICL remnant, which is bypassed by TLS. TLS past the unhooked ICL remnant depends on REV1 and Pol κ (most likely in complex with Pol ζ), which together perform both insertion opposite the colibactin-adducted base and extension of the nascent strand. A similar TLS mechanism also enables bypass of colibactin-monoadducts encountered during replication (Fig. 2.8b). Bypass of colibactin-adducts is mostly error-free, but can introduce T>A point mutations at the alkylated positions. Importantly, we observe no noticeable differences in the unhooking and bypass mechanisms for the c742-ICL and clb-ICL, supporting the use of the c742 analog for investigations into the genotoxic effects of colibactin exposure. Overall, our data provide a comprehensive model for

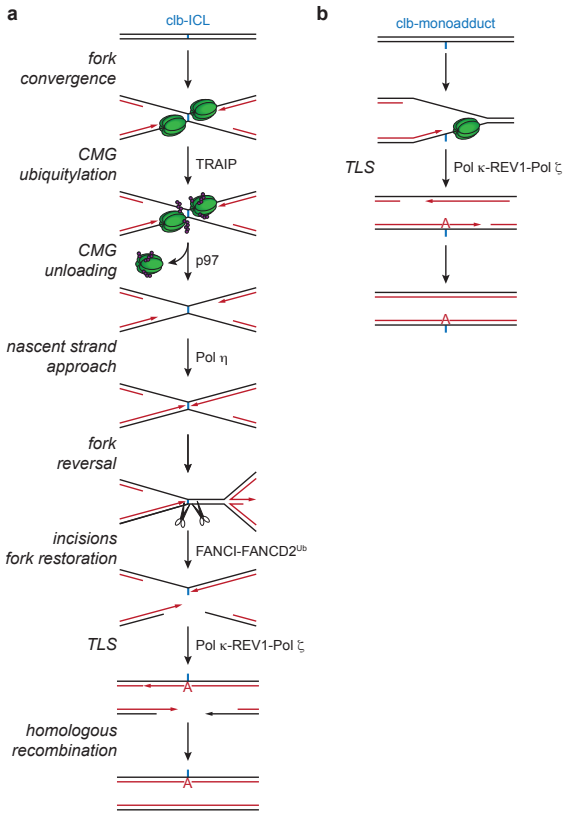


Fig. 2.8 | Model for repair of colibactin-induced DNA damage

a, Fork convergence at a colibactin-induced ICL activates CMG ubiquitylation and unloading. Pol η promotes extension of nascent leading strands up to the ICL. Following reversal of one fork, mono-ubiquitylated FANCI-FANCD2 directs nucleolytic incisions that unhook the ICL and generate a DSB. A TLS polymerase complex that includes REV1 and Pol κ enables mutagenic extension of the nascent strand past the unhooked ICL remnant, potentially introducing a T>A point mutation at the site of alkylation. The DSB intermediate is then repaired through homologous recombination.

b, A colibactin-monoadduct encountered by the replication fork is bypassed by a TLS polymerase complex that includes REV1 and Pol κ. Extension of the nascent strand past the adduct is potentially mutagenic, introducing a T>A point mutation at the site of alkylation.

replication-coupled processing of colibactin-induced DNA damage. Biallelic mutations in any of the 22 *FANC* genes linked to FA cause hypersensitivity to ICL inducing agents[13]. Moreover, the biochemical activities of FANC proteins comprise a common replication-coupled pathway that enables ICL repair. These insights have led to the presumption that

FA is caused by a failure to repair endogenous ICLs. However, a key outstanding question concerns the identities of the cognate endogenous ICLs that engage the FA pathway and presumably drive progression of bone marrow failure and cancer in FA patients. Mutations in the FA pathway strongly synergize with mutations in *ALDH2* and *ADH5*, which enable detoxification of acetaldehyde and formaldehyde, respectively[25,26], implying that unrepaired ICLs produced by endogenous reactive aldehydes cause FA. However, the specific chemical structures of aldehyde-induced ICLs that arise in human cells and drive progression of FA remain elusive. To our knowledge, the clb-ICL is the only naturally occurring ICL with both genetic and physical evidence to substantiate its formation *in vivo*[3-8]. Our data therefore demonstrate that the FA pathway repairs an ICL that is likely to arise at some point over the lifetime of most individuals.

While our data reveal that the FA pathway promotes resistance to colibactin-induced ICLs, it is unclear whether colibactin contributes to the etiology or progression of FA. We are unaware of any evidence linking *pks*$^{+}$ bacteria to bone marrow failure, and current data do not suggest that FA patients are at elevated risk for CRC, though sporadic cases have been reported[46]. Interestingly, the colibactin-associated mutational signature has been detected in squamous cell carcinomas, a cancer for which FA patients are at dramatically increased risk[47]. A recent genomic analysis of squamous cell carcinomas from FA patients did not identify colibactin associated SBSs or single nucleotide deletions[48]. However, it may be that colibactin-induced DNA damage manifests different mutational signatures in individuals lacking a functional FA pathway. In FA patients, colibactin exposure may instead produce more complex structural variants at AT-rich colibactin alkylation hotspots due to alternative processing of ICLs. The colibactin-associated mutational signature has

also been detected in cancers derived from additional tissue types[47]. Future studies are therefore warranted to determine whether FA patients are at increased risk for complications arising due to colonization by *pks*$^{+}$ bacteria.

Colibactin exhibits a striking specificity for the N3 position of dA in AT-rich motifs[10]. The three-dimensional structure of a colibactin-induced ICL implicates hydrogen bonding and electrostatic interactions within the relatively narrow minor groove of AT-rich DNA in the specificity of colibactin-ICL formation[10]. Both monoalkylation by colibactin and nucleolytic unhooking of a colibactin-induced ICL generate similar adducts comprising an N3 alkylated dA (in the latter case, the colibactin scaffold may be further linked to the short oligonucleotide generated by unhooking incisions). In general, translesion synthesis past minor groove adducts is thought to depend primarily on the polymerase activity of Pol κ. Pol κ-deficient cells are hypersensitive to genotoxins that induce minor groove adducts (e.g. Illudin S and mitomycin C), and Pol κ has been shown to efficiently extend DNA primers past N^2-dG adducts (e.g. Benzo[a]Pyrene-N^2-dG) and N3-dA adducts (e.g. 3-deaza-3-methyl-dA) *in vitro*[49,50]. Additionally, TLS past a p3d-Phen-A adduct in egg extract was recently reported to depend on Pol κ polymerase activity, but not on REV1[42]. It is therefore somewhat surprising that the bypass of colibactin-induced adducts requires REV1 but not Pol κ polymerase activity. Despite having a catalytic active site that can accommodate minor groove adducts[51], Pol κ might be inhibited by the much larger size of the colibactin adducts (clb-A molecular weight is ~900 Da). Colibactin may also make additional contacts to DNA in the minor groove that impede insertion by Pol κ. Finally, we note that previous studies have examined bypass of N3-dA adducts using stabilizing 3-

deaza-substituted dA analogs that may not completely recapitulate the effects of N3 alkylation[42].

Our data also indicate that Pol η promotes approach of nascent leading strands to within 1 nt of a colibactin-induced ICL. Based on the timing of nascent strand approach, we infer that Pol η helps to extend the nascent strands after CMG unloading and before FANCI-FANCD2-dependent incisions unhook the ICL. REV1 and Pol κ then act together to enable insertion opposite the alkylated base and subsequent extension past the lesion. Similar coordination between Pol η and REV1-Pol ζ has been observed during bypass of DNA-protein cross-links (DPCs), cisplatin-induced intrastrand cross-links, and UV-induced 6,4 photoproducts[52-54]. In these cases, Pol η normally performs insertion of a nucleotide opposite the lesion while REV1-Pol ζ performs the extension step. In the case of colibactin-ICLs, however, Pol η appears unable to perform insertion, even when REV1 is absent. It may be that the structure of the colibactin-ICL blocks insertion by Pol η. In the context of model ICL duplexes, strand displacement synthesis and insertion by Pol η was shown to be inhibited by a non-distorting ICL, resulting in an accumulation of -1 stall products[55]. The relatively non-distorting colibactin-ICL may also suppress insertion by Pol η, necessitating insertion by Pol κ-REV1-Pol ζ once unhooking has occurred[10].

Mutational signatures associated with colibactin exposure include predominantly T>C single base substitutions at AT-rich hexanucleotide sequences and single nucleotide T deletions at T-homopolymer sequences[4-6]. T>G and T>A single base substitutions are also detected, though less frequently than T>C mutations. Our data provide some insight as to the origin of these mutational signatures. Bypass of both unhooked colibactin-induced ICL remnants and monoadducts can introduce a T>A point mutation through the insertion of

dA opposite an alkylated dA. This immediately suggests that the T>A mutations observed in cells, organoids, and colonic crypts exposed to colibactin arise through Pol κ-REV1-Pol ζ-dependent TLS past colibactin adducts. This subset of colibactin-associated single-base substitutions may therefore represent a mutational signature for clb-ICL repair by the FA pathway.

How then might the more prevalent T>C mutations arise? A trivial explanation is that Pol κ-REV1-Pol ζ-dependent TLS past a colibactin adduct produces different mutations in frogs and humans due to divergence of the TLS polymerases between species. Alternatively, T>C mutations might be produced through distinct ICL repair mechanisms or through repair of distinct colibactin-induced DNA lesions. ICLs can be unhooked through replication-independent mechanisms involving mismatch repair (MMR) or transcription-coupled nucleotide excision repair (TC-NER)[45,56]. Both pathways would generate a ssDNA gap at the lesion and require TLS past an ICL remnant in the remaining strand. Interestingly, the catalytic activity of Pol κ is required for replication-independent ICL repair[45]. Colibactin-induced ICLs can decompose into DSBs through depurination and AP site cleavage, requiring DNA end joining. In such instances, the activities of alternative polymerases such as Pol λ, Pol μ, or Pol θ might be required to generate compatible DNA ends for ligation[57]. Thus, the colibactin-associated T>N mutation spectrum could reflect an amalgamation of TLS polymerase activities that act in different contexts.

When ICL unhooking or TLS is inhibited, colibactin-induced DNA lesions appear to be bypassed by a mechanism that tends to introduce larger deletions. We propose that replication forks that undergo persistent stalling at a colibactin-induced ICL undergo breakage and collapse into DSBs that are repaired through an error prone end-joining

pathway such as Pol θ-mediated end-joining (TMEJ)[58]. Similarly, Pol θ might enable microhomology-mediated gap skipping at colibactin-induced monoadducts when TLS is inhibited[59]. Consistent with this interpretation, we observe that the vast majority of the deletions detected in our sequencing data encompass at least a portion of the colibactin alkylation hotspot (Fig. 2.6b). Such deletion structures are explainable by Pol θ-dependent annealing and extension of 3' ends at sites of microhomology. It will therefore be interesting to determine whether inhibition of Pol θ and other end-joining factors influences frequencies and structures of deletions produced during colibactin-induced ICL repair. Going forward, it will also be important to examine the mutagenicity of colibactin exposure in cells carrying mutations in *FANC* genes, Pol κ, Pol ζ, Pol θ, and other TLS polymerases, and alternative ICL repair pathways. Such experiments will be critical for unambiguously mapping colibactin-associated mutational signatures to specific DNA damage tolerance mechanisms.

2.5 Methods

2.5.1 Ethics

Animal work performed at Caltech was approved by the IACUC (Protocol IA20-1797 approved 28 May 2020). The institution has an approved Animal Welfare Assurance (no. D16-00266) from the NIH Office of Laboratory Animal Welfare.

2.5.2 Statistics and Reproducibility

The experiments described in Figs. 2.1c-f, 2.2a,d, 2.3, 2.4, and 2.5 and Supplementary Figs. 2.1b-k, 2.2a,b, 2.3, 2.4, 2.5, and 2.7a-c, g-i were performed twice, and representative results

are shown. The experiments described in Fig. 2.2 b,c and Supplementary Fig. 2.2c,d were performed twice, and both sets of results are shown. The experiments described in Fig. 2.7 were performed twice, and the results for the second replicate are shown in Supplementary Fig. 2.7d-f, j-l. The experiments described in Fig. 2.6 were performed three times, and the results of the second and third replicates are shown in Supplementary Fig. 2.6d-f. The experiments described in Supplementary Fig. 2.6a-c were performed three times, and representative results are shown.

2.5.3 Oligonucleotides

Oligonucleotides used for preparation of NGS libraries, cross-linked inserts, and sequencing ladders were ordered from Integrated DNA Technologies, and their sequences can be found in Supplementary Table 1.

2.5.4 Preparation of Xenopus egg extracts

Preparation of high-speed supernatant (HSS) and nucleoplasmic extracts (NPE) from *Xenopus laevis* eggs was performed as described previously[60]. Briefly, HSS was prepared from eggs collected from six laboratory bred wild-type adult female *X. laevis* (aged >2 years). Eggs were dejellied in 1 L of 2.2% (w/v) cysteine, pH 7.7, washed with 2 L 0.5X Marc's Modified Ringer's solution (2.5 mM HEPES-KOH [pH 7.8], 50 mM NaCl, 1 mM KCl, 0.25 mM $MgSO_4$, 1.25 mM $CaCl_2$, 0.05 mM EDTA), and washed with 1 L of Egg Lysis Buffer (ELB) (10 mM HEPES-KOH [pH 7.7], 50 mM KCl, 2.5 mM $MgCl_2$, 250 mM sucrose, 1 mM dithiothreitol [DTT], and 50 µg/mL cycloheximide). Eggs were then packed in 14-mL round-bottom Falcon tubes at 200 x g using a Sorvall ST8 swinging bucket rotor. Eggs were supplemented with 5 µg/mL aprotinin, 5 µg/mL leupeptin, and 2.5 µg/mL

cytochalasin B and then crushed by centrifugation at 20,000 x g for 20 min at 4 °C in a TH13-6x50 rotor using a Sorvall Lynx 4000 centrifuge. The low-speed supernatant (LSS) was collected by removing the soluble extract layer and supplemented with 50 µg/mL cycloheximide, 1 mM DTT, 10 µg/mL aprotinin, 10 µg/mL leupeptin, and 5 µg/mL cytochalasin B. This extract was then spun in thin-walled ultracentrifuge tubes at 260,000 x g for 90 min at 2 °C in a TLS-55 rotor using a tabletop ultracentrifuge. Lipids were aspirated off the top layer, and HSS was harvested, aliquoted, snap frozen in liquid nitrogen, and stored at -80 °C. NPE preparation also began with extracting LSS, except eggs were collected from 20 laboratory bred wild-type female *X. laevis* (aged >2 years), and the volumes used to dejelly and wash the eggs were doubled (2 L of 2.2% cysteine, 4 L of 0.5X Marc's Modified Ringer's solution, and 2 L of ELB). LSS was supplemented with 50 µg/mL cycloheximide, 1 mM DTT, 10 µg/mL aprotinin, 10 µg/mL leupeptin, 5 µg/mL cytochalasin B, and 3.3 µg/mL nocodazole. The LSS was then spun at 20,000 x g for 15 min at 4 °C in a TH13-6x50 rotor using a Sorvall Lynx 4000 centrifuge. The top, lipid layer was removed, and the cytoplasm was transferred to a 50-mL conical tube. ATP regenerating mix (2 mM ATP, 20 mM phosphocreatine and 5 µg/mL creatine phosphokinase) was added to the extract. Nuclear assembly reactions were initiated by adding demembranated *X. laevis* sperm chromatin[60] to a final concentration of 6,600 units/µL. After 75-90 min incubation, the nuclear assembly reactions were centrifuged for 3 min at 17,000 x g at 4 °C in a TH13-6x50 rotor using a Sorvall Lynx 4000 centrifuge. The top, nuclear layer was then harvested and spun at 260,000 x g for 30 min at 2 °C in a TLS-55 rotor using a tabletop ultracentrifuge. Finally, lipids were aspirated off the top

layer, and NPE was harvested, aliquoted, snap frozen in liquid nitrogen, and stored at -80 °C.

2.5.5 Protein expression and purification

Recombinant *Xenopus* FANCI-FANCD2 complex was expressed and purified essentially as previously described[37]. Baculoviruses expressing the FLAG-FANCI or Strep-FANCD2 proteins were prepared using the Bac-to-Bac Baculovirus Expression System according to the manufacturer's protocols. 500 mL Sf9 insect cell suspension cultures were co-infected with viruses expressing FLAG-FANCI and either Strep-FANCD2 or FLAG-FANCI and Strep-FANCD2^{K562R} and incubated for 72 hours at 27 °C. Cells were collected by spinning at 500 x g for 10 minutes at 4 °C and washed with 1X PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4). Cell pellets were frozen in liquid nitrogen and stored at -80 °C. All subsequent steps were performed at 4 °C. Cell pellets were thawed and resuspended in 10 mL lysis buffer (50 mM Tris-HCL pH 8.0, 100 mM NaCl, 0.5 mM PMSF, 1 mM EDTA, 1 tablet EDTA-free protease inhibitor cocktail [Roche] in 50 mL total volume). The cells were passed through a 21G needle 3 times and lysed by sonication in a Misonix S-4000 sonicator with a 0.25-inch microtip at 40% output for 1 minute (1 second on, 1 second off) for 4 rounds with 20 seconds between rounds. After centrifugation at 40,000 x g for 40 minutes, the soluble fraction was collected and filtered through a 0.45 µm syringe filter. The filtrate was incubated for 2 hours with 200 µL anti-FLAG M2 affinity resin (Sigma) that was prewashed with lysis buffer. During the incubation, a 5 mL polypropylene column was packed with 0.5 mL Strep-Tactin XT 4Flow resin (IBA Life Sciences), overlaid with 0.5 mL Strep wash buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA), and stored at 4 °C until use. After the incubation, the anti-FLAG

beads with bound protein complex were washed once with 10 mL lysis buffer and 4 times with 10 mL FLAG wash buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 μg/mL aprotinin/leupeptin) and collected by spinning at 1,000 x g for 2 minutes. The resin was transferred to a 1 mL polypropylene column, and the FANCI-FANCD2 complex was eluted with FLAG wash buffer containing 100 μg/mL 3X FLAG peptide (Sigma) and 5% glycerol. The resin was incubated for 15 minutes in 200 μL elution buffer before each fraction was collected, with the exception of the first fraction, for which the incubation time was 40 minutes. Fractions containing FANCI-FANCD2 were pooled and applied to the Strep-Tactin column, which was pre-washed with 1 mL (2 column volumes) of Strep wash buffer. The flow-through was collected and re-applied to the column, which was then incubated for 30 minutes. The column was subsequently washed 5 times with 0.5 mL Strep wash buffer, and the protein complex was eluted with Strep wash buffer containing 50 mM biotin in 0.5 CV fractions. Fractions containing FANCI-FANCD2 were pooled and dialyzed against freezing buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 20% glycerol) overnight. The protein complex was concentrated to ~5 μM with a 0.5 mL Amicon filter unit (10 kDa MWCO) and stored at -80 °C. Where indicated, rFANCI-D2 or rFANCI-D2^{K562R} was added to replication reactions to a final concentration of 375 nM in extracts.

Recombinant Pol η was prepared using Promega TNT Quick Coupled Transcription/Translation system according to the manufacturer's instructions. 2 μg pCMV-Sport6-POLH plasmid was mixed with 80 μL TNT master mix supplemented with 20 μM methionine in 100 μL total volume. The reaction mixture was incubated at 30 °C for 90 minutes. The reaction was then concentrated to ~25 μL by centrifugation at 14,000 x g for 20 minutes in an Amicon 0.5 mL filter unit (30 kDa MWCO). Aliquots were then

frozen in liquid nitrogen and stored at -80 °C. 0.25 volumes of the rPol η in vitro transcription/translation mixture were used to supplement 1 volume of NPE.

To prepare recombinant Pol κ, Arctic Express cells were transformed with pET28-6xHis-Pol κ and incubated at 30 °C in Luria broth supplemented with 20 μg/mL Gentamicin, and 50 μg/mL kanamycin. Once the OD_{600} reached 0.6, cells were incubated for 20 hours at 11.5 °C in the presence of 1 mM isopropyl b-D-1-thiogalactopyranoside. Cell pellets were lysed in 50 mM Tris-HCl (pH 7.6), 300 mM NaCl, 0.1% Triton X-100, 10 mM imidazole, 2 mM $MgCl_2$, 1 mM DTT, 1 mg/mL lysozyme at 4 °C for 45 min. Cells were then sonicated using a Misonix S-4000 sonicator with 0.25-inch microtip for a total of 2 minutes at 40% amplitude. The lysate was clarified by spinning at 13,000 x g for 60 minutes at 4 °C and then incubated with Ni-NTA resin at 4 °C for 5 hours. The resin was then washed with 50 mM Tris-HCl (pH 7.6), 400 mM NaCl, 0.1% Triton X-100, 40 mM imidazole, 2 mM $MgCl_2$, 1 mM DTT, incubated in 50 mM Tris-HCl (pH 7.6), 400 mM NaCl, 150 mM KCl, 20% glycerol, 0.1% Triton X-100, 10 mM imidazole, 10 mM $MgCl_2$, 10 mM BME, 5 mM ATP at 4 °C for 3 hours, and washed with 50 mM Tris-HCl (pH 7.6), 400 mM NaCl, 0.1% Triton X-100, 40 mM imidazole, 2 mM $MgCl_2$, 1 mM DTT. Pol κ was eluted with 50 mM Tris-HCl (pH 7.6), 300 mM NaCl, 300 mM imidazole, 2 mM $MgCl_2$, 1 mM DTT, 10% glycerol and dialyzed into 50 mM Tris-HCl (pH 7.6), 300 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT, 10% glycerol at 4 °C. Aliquots were frozen in liquid nitrogen and stored at -80 °C. Where indicated rPol κ was added to NPE at a concentration of 75 nM.

2.5.6 Immunodepletions

Immunodepletions using antibodies raised against FANCD2, the N- or C-terminus of REV1, Pol η, and Pol κ were performed as previously described[23,37,42,52]. Briefly, Protein A Sepharose Fast Flow (Cytiva) resin was washed with 1X PBS and incubated with an appropriate volume of antibody overnight at 4 °C (3.6 volumes for FANCD2, 1 volume for REV1, and 5 volumes for Pol κ and Pol η). The next day, the beads were washed twice with 1X PBS, once with ELB mix (2.5 mM $MgCl_2$, 50 mM KCl, 10 mM HEPES, 250 mM sucrose), twice with ELB mix supplemented with 0.5 M NaCl, and 3 times with ELB mix. For FANCD2 depletion, 2 rounds of depletion were performed by adding 7.14 volumes of egg extracts (HSS and NPE separately) to 1 volume of beads and incubating on a rotating wheel at room temperature for 20 mins per round. For REV1 depletion, 5 volumes of egg extracts were added to 1 volume of beads and incubated at 4 °C for 60 minutes per round; 2 rounds of depletion were performed with REV1-N beads, followed by one round with REV1-C beads. For Pol κ and Pol η depletions, 5 volumes of egg extracts were added to 1 volume of beads and incubated at room temperature for 15 minutes per round; one round of depletion was performed for HSS and 3 rounds for NPE. Between rounds of depletion, the samples were spun at 622 x g for 30 seconds in an S-24-11-AT rotor in an Eppendorf 5430R centrifuge, and extract supernatants were collected.

2.5.7 Preparation of oligonucleotide duplexes with site-specific interstrand cross-links

Site specific cross-links were prepared as previously described[15,61]. To generate the c742-ICL containing oligonucleotide duplex, complementary oligonucleotides (c742-ICL top and c742-ICL bottom; 151 μM each) were mixed in 10 mM $NaPO_4$ buffer (pH 6.0). The mixture was heated to 95 °C for 5 min and then cooled at 1 °C/min to 18 °C. Cross-linking was performed by incubating 40 μM annealed duplex with 30 μM c742 colibactin analog

in 10 mM $NaPO_4$ buffer (pH 6.0) at 37 °C for 4 hours. The mixture was adjusted to 10 mM NaOH, 100 mM NaCl and the cross-linked oligonucleotide duplex was then purified over a Mono Q 5/50 GL column using a gradient from 550 mM to 700 mM NaCl in 10 mM NaOH over 40 column volumes. Fractions containing the cross-linked duplex were combined and buffer exchanged into 10 mM NaOH, 100 mM NaCl with a 4 mL Amicon filter unit (10 kDa MWCO). A second round of purification was performed over a Mono Q 5/50 GL column using a gradient from 550 mM to 700 mM NaCl in 10 mM NaOH over 40 column volumes. Fractions containing the cross-linked duplex were combined and buffer exchanged into 10 mM Tris-HCl (pH 8.8) with a 4 mL Amicon filter unit (10 kDa MWCO). Aliquots were frozen in liquid nitrogen and stored at -20 °C.

To generate the clb-ICL containing oligonucleotide duplex, complementary oligonucleotides (clb-ICL top and clb-ICL bottom; 20 μM) were mixed 1:1 in annealing buffer (10 mM Tris-HCl [pH 8.0], 50 mM NaCl). The mixture was heated to 95 °C and then cooled at 1.5 °C/min to 20 °C. Annealed samples were stored at 4 °C and warmed to 25 °C prior to use. Overnight cultures of *E. coli* BW25113 + BAC-*pks* grown in LB were back-diluted 1:100 into fresh M9-CAS and shaken at 37 °C to an OD_{600} of 0.20. Annealed oligonucleotide samples were added directly into the bacteria culture at a dilution factor of 1:40, and the mixture was dispensed into 96-well plates (200 μL per well; Corning). The plates were further incubated at 37 °C without shaking for 5 hours. The contents in the 96-well plates were then recombined and spun down at 4,000 x g for 20 mins, and the supernatant was carefully decanted and filtered through a 0.22 μM PES filter (Corning), flash frozen in liquid nitrogen, and lyophilized to completely dryness. Dried samples were reconstituted in Oligo Binding Buffer (Zymo) and purified using the Oligo Clean &

Concentrator™ Kit (Zymo) following the manufacturer's protocol. The cross-linked oligonucleotide duplex was purified over a Mono Q 5/50 GL column using a gradient from 550 mM to 700 mM NaCl in 10 mM NaOH over 40 column volumes. Fractions containing the cross-linked duplex were combined and buffer exchanged into 10 mM NaOH, 100 mM NaCl with a 4 mL Amicon filter unit (10 kDa MWCO). The cross-linked duplex was further purified on a 10% acrylamide/bis (19:1), 1X Tris-borate-EDTA (TBE; 89 M Tris, 89 mM boric acid, 2 mM EDTA), 8 M urea gel. The gel was stained with SYBR Gold, and the cross-linked duplex was visualized with a Blue-Light Transilluminator. The clb-ICL duplex was excised, eluted from crushed gel slices into 10 mM Tris-HCl (pH 8.8) on a rotating wheel at 4 °C overnight, extracted with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0), and precipitated in ethanol. The purified cross-linked duplex was dissolved in 10 mM Tris-HCl (pH 8.8) and stored at -80 °C.

To generate the Pt-ICL containing oligonucleotide duplex, 1 mM cisplatin was converted to activated monoaquamonochloro cisplatin by incubation in 10 mM $NaClO_4$ (pH 5.2), 0.95 mM $AgNO_3$ for 24 hours at 37 °C in the dark. Cisplatin monoadduct was then generated by incubating 0.125 mM Pt-ICL top oligonucleotide in 5.63 mM $NaClO_4$ (pH 5.2), 0.375 mM monoaquamonochloro cisplatin (in activation mixture) for 12 min at 37 °C. The reaction was quenched by addition of NaCl to 0.1 M. The monoadducted oligonucleotide was then purified over a Mono Q 5/50 GL column using a gradient from 370 mM to 470 mM NaCl in 10 mM NaOH over 40 column volumes. Fractions containing the monoadducted oligonucleotide were pooled and adjusted to 2 mM $MgCl_2$. 1.05 molar equivalents of Pt-ICL bottom oligonucleotide were added, and buffer exchange was performed with 100 mM $NaClO_4$ using an Amicon Ultra-15 3K filter unit at 4 °C. The

oligonucleotides in 100 mM $NaClO_4$ were allowed to cross-link by incubation at 37 °C for 48 hours. The cross-linked oligonucleotide duplex was then purified on a Mono Q 5/50 GL column using a gradient from 550 mM to 700 mM NaCl in 10 mM NaOH over 40 column volumes. Fractions containing the Pt-ICL duplex were pooled and concentrated using an Amicon Ultra-15 3K filter unit at 4 °C. The Pt-ICL duplex was further purified on a 20% acrylamide/bis (19:1), 1X TBE, 8 M urea gel. The gel was stained with SYBR Gold, and the cross-linked duplex was visualized with a Blue-Light Transilluminator. The Pt-ICL duplex was excised, eluted from crushed gel slices into TE (pH 8.0), extracted with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0), and ethanol precipitated. The purified cross-linked duplex was dissolved in 10 mM Tris-HCl (pH 7.4), 10 mM $NaClO_4$ and stored at -80 °C.

2.5.8 Preparation of plasmids containing cross-links (pICL)

Plasmids containing ICLs were prepared as described previously[15,16,61]. Briefly, the backbone plasmid (containing 48 lacO repeats) was digested with BbsI in NEBuffer r2.1 (10 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 100 µg/mL recombinant albumin; pH 7.9) for 24 hours at 37 °C followed by extraction with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8) and ethanol precipitation. The linearized plasmid was dissolved in TE (pH 8.0) and purified over a HiLoad 16/60 Superdex 200 column using isocratic flow of TE (pH 8.0). Fractions containing the digested plasmid were pooled, ethanol precipitated, and dissolved in 10 mM Tris-HCl (pH 8.5). $pICL^{c742}$ was prepared by incubating 0.5 nM BbsI-digested plasmid with 1.5 nM c742-ICL duplex and 0.4 U/µL NEB T4 DNA ligase in 1X ligase buffer (50 mM Tris-HCl [pH 8.0], 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT) at 4 °C for 6 hours. The ligation reaction was then concentrated using a 15 mL Amicon filter unit

(10 kDa MWCO), extracted with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8), and then buffer exchanged into 10 mM Tris-HCl (pH 8.8), 50 mM NaCl, 1 mM $MgCl_2$ with a 4 mL Amicon filter unit (10 kDa MWCO). CsCl was added to the DNA mixture to a homogenous solution density of 1.58 g/mL and ethidium bromide was added to 50 µg/mL. The DNA was then transferred to a 3.5 mL Quick-Seal tube (Beckman Coulter) and spun for 13 hours at 4 °C in a TLA-100.3 rotor at 279,682.5 x g. The covalently closed circular plasmid was collected and extracted 6 times with an equal volume of saturated n-butanol to remove ethidium bromide. The DNA was then buffer exchanged into 10 mM Tris-HCl (pH 8.8) with a 4 mL Amicon filter unit (10 kDa MWCO). Aliquots of $pICL^{c742}$ were frozen in liquid nitrogen and stored at -80 °C.

$pICL^{clb}$ was prepared by incubating 0.5 nM BbsI digested plasmid with 0.75 nM clb-ICL duplex and 0.4 U/µL NEB T4 DNA ligase in 1X ligase buffer (50 mM Tris-HCl [pH 8.0], 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT) at 4 °C for 16 hours. The ligation reaction was then concentrated using a 15 mL Amicon filter unit (30 kDa MWCO), extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0) and once with chloroform, ethanol precipitated, and resuspended in 10 mM Tris-HCl (pH 8.8), 50 mM NaCl, 1 mM $MgCl_2$. For every 1 mL of the concentrated DNA mixture, 1 g CsCl was added to the mixture. The DNA was transferred to a 5.1 mL Quick-Seal tube (Beckman Coulter) and supplemented with ethidium bromide to 50 µg/mL. The tube was filled to the top with a CsCl-buffer solution consisting of 1 g of CsCl per 1 mL of 10 mM Tris-HCl (pH 8.8), 50 mM NaCl, 1 mM $MgCl_2$ buffer. The mixture was spun for 16 hours at 4 °C in an NVT-90 rotor at 249,178 x g. The covalently closed circular plasmid was collected and extracted 6 times with an equal volume of saturated n-butanol to remove ethidium bromide. The DNA was

then buffer exchanged into 10 mM Tris-HCl (pH 8.8), 50 mM NaCl, 1 mM $MgCl_2$ with a 4 mL Amicon filter unit (30 kDa MWCO). Aliquots of $pICL^{clb}$ were frozen in liquid nitrogen and stored at -80 °C.

$pICL^{Pt}$ and $pICL^{AP}$ were prepared by incubating 0.5 nM BbsI digested plasmid with 0.75 nM Pt-ICL or AP-ICL duplex and 0.4 U/µL NEB T4 DNA ligase in 1X ligase buffer (50 mM Tris-HCl [pH 7.5], 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT) at 4 °C for 16 hours. The ligation reaction was then concentrated using a 15 mL Amicon filter unit (30 kDa MWCO), extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0) and once with chloroform, ethanol precipitated, and resuspended in TE. For every 1 mL of the concentrated DNA mixture, 1 g CsCl was added to the mixture. The DNA was transferred to a 5.1 mL Quick-Seal tube (Beckman Coulter) and supplemented with ethidium bromide to 50 µg/mL. The tube was filled to the top with a CsCl-TE solution consisting of 1 g of CsCl per 1 mL of TE. The mixture was spun for 23 hours at 20°C in an NVT-90 rotor at 389,340 x g. The covalently closed circular plasmid was collected and extracted 6 times with an equal volume of saturated n-butanol to remove ethidium bromide. The DNA was then buffer exchanged into 10 mM Tris-HCl (pH 7.5) for $pICL^{AP}$ and 10 mM Tris-HCl (pH 7.4), 10 mM $NaClO_4$ for $pICL^{Pt}$ with a 4 mL Amicon filter unit (30 kDa MWCO). Aliquots of both plasmids were frozen in liquid nitrogen and stored at -80 °C.

2.5.9 Replication reactions

Plasmid replication reactions in *Xenopus* egg extracts were performed essentially as previously described[62]. Briefly, HSS was thawed and supplemented with 20 mM phosphocreatine, 2 mM ATP, 5 µg/mL creatine phosphokinase, and 3 µg/mL nocodazole.

Plasmids were added to the HSS mix to a concentration of 7.5 ng/μL and incubated at room temperature for 30 mins to enable licensing. For assays involving LacR acting as a replication block, plasmids were pre-incubated with 14 μM biotinylated LacI for 1 hour at room temperature before being added to HSS. Where indicated, the licensing mix was supplemented with 167-333 nM 3000 Ci/mmoL [α-^{32}P]dCTP and/or 400 nM rGeminin. NPE mix was made with 50% NPE, 20 mM phosphocreatine, 2 mM ATP, 5 μg/mL creatine phosphokinase, and 4 mM DTT in ELB. Where indicated, NPE mix was supplemented with 50 μM PHA-767491 HCl or 300 μM NMS-873. Replication was initiated by addition of 2 volumes of NPE mix to 1 volume of licensing mix. Where indicated, 10 mM IPTG was added 15 minutes into the replication time course to dissociate LacI from the plasmid. For samples analyzed by APE1 digestion, HSS and NPE mixes were combined before adding plasmid (i.e. replication origins were not licensed in HSS mix). For analysis on native agarose gels, 1 μL samples of the reaction were removed at indicated time points and quenched by addition to 6 μL of replication stop mix (80 mM Tris-HCl [pH 8.0], 8 mM EDTA, 0.13% phosphoric acid, 10% ficoll, 5% SDS, 0.1% bromophenol blue). After all time points were quenched, samples were digested with 1.25 mg/mL proteinase K for 60 mins at 37 °C, and replication intermediates and products were resolved on 0.8% agarose, 1X TBE gels. Gels were dried and visualized by phosphorimaging. Autoradiograms were imaged with a Typhoon FLA 9500 (GE Healthcare) using the FujiFilm FLA 9500 user interface v.1.1 and analyzed using Image Lab v.6.4.0. For Southern blots and nascent strand analysis on sequencing gels, 4 μL samples of the reaction were removed at indicated time points and quenched by addition to 40 μL of clear replication stop mix (50 mM Tris-HCl [pH 8.8], 0.5% SDS, 25 mM EDTA). After all time

points were quenched, samples were digested with 200 µg/mL RNAse A for 30 mins at 37 °C and then with 1 mg/mL proteinase K for 60 mins at 37 °C. Samples were adjusted to 200 µL with 10 mM Tris-HCl [pH 8.8], extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0) and once with chloroform, and ethanol precipitated. DNA was resuspended in 10 µL of 10 mM Tris-HCl (pH 8.8), 50 mM NaCl, 1 mM $MgCl_2$ (for c742-ICL and clb-ICL plasmids) or 10 mM Tris-HCl pH 8.5 (for all other plasmids) and stored at -20 °C (or -80 °C for plasmids intended for unhooking assays). For samples analyzed by APE1 digestion, 3 µL of recovered DNA were incubated with 1 U APE1 in 1X NEBuffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM DTT; pH 7.9) in a total volume of 10 µL at 37 °C for 60 mins. Samples were subsequently run on a 0.8% agarose, 1X TBE, 0.5 µg/mL ethidium bromide gel and visualized with a UV transilluminator.

2.5.10 Southern blotting

Southern blotting was performed essentially as previously described[37]. Briefly, samples were replicated as described above, and 3 µL of recovered DNA were incubated with 5 U HincII in 1X NEBuffer rCutsmart (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 µg/mL recombinant albumin; pH 7.9) in a total volume of 5 µL at 37 °C for 60 mins. Following digestion, 5 µL of 50 mM Tris-HCl (pH 8.8), 50 mM NaCl, 10 mM EDTA and 2 µL of 6X alkaline loading buffer (300 mM NaOH, 6 mM EDTA, 18% ficoll, 0.25% xylene cyanol, 0.15% bromocresol green) was added to each sample. The DNA was resolved on a denaturing 0.8% agarose, 50 mM NaOH, 1 mM EDTA gel for 20 hours at 1.25 V/cm. After running was complete, the gel was incubated in 0.25 M HCl for 20 mins and in 0.4 M NaOH, 3 M NaCl for 60 minutes. Downward capillary transfer onto

an Invitrogen BrightStar-Plus positively charged nylon membrane was performed overnight in 0.4 M NaOH, 1.5 M NaCl. The membrane was subsequently washed with 4X SSC buffer (0.6 M NaCl, 60 mM sodium citrate; pH 7.0) for 5 mins, and the DNA was cross-linked to the membrane via 265 nm UV irradiation (120,000 μJ cm^{-2}) with a UVP HybriLinker Oven. The membrane was then incubated with 15 mL ULTRAhyb Ultrasensitive Hybridization Buffer (Invitrogen) for 1-2 hours at 42 °C. During this prehybridization, a radioactive probe was made using 50 ng of linearized pCtrl and the Prime-a-Gene Labeling System kit (Promega) following the manufacturer's instructions. The probe was added to the hybridization tube and incubated with the membrane at 42 °C for 18 hours. After hybridization, the membrane was washed twice with 1X SSC, 0.7% SDS at 42 °C for 10-20 minutes per wash and twice with 0.1X SSC, 0.7% SDS that was pre-heated to 75 °C for 10-20 minutes per wash. The membrane was dried briefly on blotting paper and visualized by phosphorimaging as described above.

2.5.11 Nascent strand analysis

Nascent strand analysis was performed essentially as previously described[36]. Briefly, samples were replicated as described above, and 2 μL of recovered DNA were incubated with 2 U AflIII and 4 U of either EcoRI or BamHI in 50 mM Tris-HCl (pH 8), 100 mM NaCl, 10 mM $MgCl_2$ in a total volume of 5 μL at 37 °C for 3 hours. Reactions were stopped by adding 5 μL of Invitrogen Gel Loading Buffer II. Primers for sequencing ladders were radiolabeled by incubating 200 nM sequencing primer with 200 nM [γ-^{32}P]ATP and 20 U T4 polynucleotide kinase (NEB) in 1X T4 PNK reaction buffer (70 mM Tris-HCl, 10 mM $MgCl_2$, 5 mM DTT; pH 7.6) in a total volume of 50 μL for 30 minutes at 37 °C. T4 PNK was inactivated by heating at 95 °C for 2 minutes, and the unincorporated radionucleotides

were removed using a Micro Bio-Spin column (Bio-Rad) per the manufacturer's instructions. Sequencing ladders were made using the Thermo Sequenase Cycle Sequencing Kit: 0.5 pmol of radiolabeled primer were mixed with 175 ng of pCtrl, 2 µL of reaction buffer, and 2 µL of sequenase in a total volume of 17.5. 4 µL of this reaction mix were then added to 4 µL of ddA, ddG, ddC, or ddT termination mixes, and the reactions were incubated on a PCR block at 95 °C for 3 minutes, followed by 50 cycles of 95 °C for 30 seconds, 60 °C for 60 seconds, and 72 °C for 90 seconds. 8 µL of Invitrogen Gel Loading Buffer II were added to each reaction, and the sequencing ladders were stored at -20 °C until use. Immediately before loading, all samples were heated at 75 °C for 5 minutes and snap-cooled in an ice water bath. DNA was resolved on a 7% acrylamide/bis (19:1), 8 M urea, 0.8X GTG buffer (71 mM Tris-HCl, 23 mM taurine, 0.4 mM EDTA) gel. Sequencing gels were dried and visualized by phosphorimaging as described above.

2.5.12 Plasmid pulldown and sample preparation for mass spectrometry

Plasmid pulldowns for the purpose of mass spectrometry were performed essentially as previously described[40]. Briefly, 6 µL of M-280 Streptavidin Dynabeads (Invitrogen) per time point were washed 3 times with 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.02% Tween-20, and 1 mM EDTA (pH 8.0) and subsequently incubated in the same buffer supplemented with 2 pmol of biotinylated LacI per µL of beads for 40 minutes at room temperature on a rotator wheel. The beads were washed 4 times with pulldown buffer (2.5 mM $MgCl_2$, 50 mM KCl, 10 mM HEPES, 250 mM sucrose, 0.25 mg/mL BSA, 0.02% Tween-20), divided into 40 µL aliquots, and stored on ice. Replication reactions were performed as described above, except extracts were supplemented with 40 µg/mL RNAse A and 10 U/µL RNAse T1, and plasmids were replicated at a final concentration of 5 ng/µL

in extracts. At the indicated times, 8 µL of each reaction were quenched into 40 µL of Dynabeads (with 4 replicates for each time point) and immediately incubated on a rotator wheel for 30 minutes at 4 °C to enable binding of the plasmid's *lacO* array to the LacI-coated beads. After the incubation, the beads were washed twice at 4 °C with 2.5 mM $MgCl_2$, 50 mM KCl, 10 mM HEPES, 250 mM sucrose, 0.03% Tween-20 and once with 2.5 mM $MgCl_2$, 50 mM KCl, 10 mM HEPES, 250 mM sucrose. For each time point, the beads for one of the replicates were suspended in 10 µL of 1X Laemmli loading buffer and analyzed by immunoblotting. Beads for the other 3 replicates were suspended in 50 µL of 50 mM HEPES-KOH (pH 8.0), transferred to fresh microtubes, and stored on ice. The buffer was aspirated from the beads, and they were resuspended in 50 µL of 8 M urea in 50 mM HEPES-KOH (pH 8.0). Subsequent steps were performed at room temperature unless otherwise indicated. Sample were incubated with 20 mM DTT for 15 minutes, supplemented with chloroacetamide to 58 mM, and incubated at 1,200 rpm on a Thermomixer for 40 minutes (covered in foil to protect from light). Samples were then incubated with 1.5 µg Lys-C at 1,400 rpm on a Thermomixer for 150 minutes. Following the incubation, supernatants containing the eluted peptides were separated from the beads on a magnetic rack and transferred to fresh tubes (~60 µL of supernatant per sample). Samples were adjusted to 200 µL with 50 mM ammonium bicarbonate and incubated with 2.5 µg of trypsin at 30 °C overnight. Digestions were terminated by adding 20 µL of 10% TFA (trifluoroacetic acid) and adjusted to 0.5 M NaCl by adding 74 µL of 2 M NaCl. Stage tips were prepared by dispensing two C18 disks into 200 µL pipette tips with a blunt needle and spinning at 500 x g for 3 mins to pack the disks. All subsequent stage tip spins were also performed at 500 x g for 3 mins. Stage tips were activated by spinning with 100 µL

50% acetonitrile (ACN) and equilibrated by spinning with 100 µL 0.1% TFA. Samples were spun at 8,000 x g for 10 mins to pellet any undigested DNA debris. 85 µL of supernatant were applied to stage tips and spun; this step was repeated two more times, leaving ~40 µL of each sample behind to avoid disturbing the pellet. Stage tips were washed twice with 100 µL of 0.1% TFA, and peptides were eluted twice with 50 µL of 50% ACN, 0.1% TFA. Samples were then dried in a SpeedVac, reconstituted in 10 µL of 2% ACN/0.2% FA (formic acid) solution, sonicated for 5 mins, and centrifuged at 16,000 x g for 5 mins. 4 µL of each sample were transferred to autosampler vials. LC-MS/MS analysis was performed with an EASY-nLC 1200 (ThermoFisher Scientific, San Jose, CA) coupled to a Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer (ThermoFisher Scientific, San Jose, CA). Peptides were separated on an Aurora UHPLC Column (25 cm x 75 µm, 1.7 µm C18, AUR3-25075C18, Ion Opticks) with a flow rate of 0.35 µL/min for a total duration of 75 mins and ionized at 1.6 kV in the positive ion mode. The gradient was composed of 6% solvent B (3.5 mins), 6-25% B (42 mins), 25-40% B (14.5 mins), and 40-98% B (15 mins); solvent A: 2% ACN and 0.2% formic acid in water; solvent B: 80% ACN and 0.2% formic acid. MS1 scans were acquired at a resolution of 60,000 from 375 to 1500 m/z, AGC target 3E6, and maximum injection time 15 ms. The 12 most abundant ions in MS2 scans were acquired at a resolution of 30,000, AGC target 1E5, maximum injection time 60 ms, and a normalized collision energy of 28. Dynamic exclusion was set to 30 s, and ions with charges of +1, +7, +8 and >+8 were excluded. The temperature of the ion transfer tube was 275 °C, and the S-lens RF level was set to 60.

2.5.13 Proteomic data processing

Label-free quantification (LFQ) of protein abundances was performed with the MaxLFQ algorithm implemented in MaxQuant (MQ v2.3.1.0)[63-65]. The analysis used default settings with match between runs (MBR) functionality enabled. RAW data were searched against an *in silico* tryptic digested protein reference database based on the 10.1 version of the *Xenopus laevis* genome downloaded from Xenbase (https://wuehr.scholar.princeton.edu/sites/g/files/toruqf2761/files/xenla_10_1_xenbase.fasta_.txt). The maximum missed cleavages were set to 2. Dynamic modifications were set to oxidation on methionine (M, +15.995 Da) and protein N-terminal acetylation (+42.011 Da). Carbamidomethylation on cysteine residues (C, +57.021 Da) was set as a fixed modification. Protein quantification was performed using the built-in label-free quantification (LFQ) algorithm. The data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD055646[66].

The Perseus software platform (v.2.1.2.0) was used for data filtering and normalization[67]. Potential contaminant proteins, reverse database protein hits, proteins identified only by modification site, and proteins that were not quantified in at least one replicate group were removed from the analysis. LFQ intensities were log2 transformed, and samples were median normalized. Missing abundance values were imputed from a downshifted normal distribution with default parameters defined in Perseus. For heatmap visualization, normalized abundances were averaged within biological replicate groups and standardized. The R package *limma* was used for statistical analysis of protein abundances[68]. Differential expression p-values from empirical Bayes moderated t-testing of the log2-transformed protein ratios were adjusted for multiple hypothesis testing with the Benjamini-Hochberg procedure[69].

2.5.14 Immunoblotting

Samples in 1X Laemmli loading buffer (generally corresponding to 1-2 μL of replication reaction or 9 μL of a plasmid pulldown sample) were resolved on 10% or 4-15% acrylamide Mini-PROTEAN or Criterion TGX precast gels (Bio-Rad) and transferred to polyvinyl difluoride (PVDF) membranes (Thermo Fisher). Membranes were blocked with 5% nonfat milk in 1X PBS buffer with 0.05% (v/v) Tween 20 (PBST) for 1 hour at room temperature, rinsed 4 times with 1X PBST, and incubated with primary antibodies diluted to an appropriate concentration in 1X PBST at 4 °C overnight with shaking. Primary antibody dilutions were as follows: Rabbit polyclonal anti-MCM7 (*Xenopus*) Pocono Rabbit Farm and Laboratory rabbit 456, 1:12,000; Rabbit polyclonal anti-CDC45 (*Xenopus*) Pocono Rabbit Farm and Laboratory rabbit 534, 1:20,000; Rabbit polyclonal anti-H3 (human) Cell Signaling Technology catalog no. 9715S, 1:500; Rabbit polyclonal anti-REV1 (C terminus; *Xenopus*) Pocono Rabbit Farm and Laboratory rabbit 714, 1:5,000; Rabbit polyclonal anti-FANCI (*Xenopus*) Pocono Rabbit Farm and Laboratory rabbit 26864, 1:5,000; and Rabbit polyclonal anti-FANCD2 (*Xenopus*) Pocono Rabbit Farm and Laboratory rabbit 20019, 1:5,000; Rabbit polyclonal anti-Pol η New England Peptide rabbit 2991, 1:5,000, and Rabbit polyclonal anti-Pol κ New England Peptide rabbit 3924, 1:5000[52]. The next day, the membranes were washed with 1X PBST 3 times for 10 minutes each at room temperature, then incubated for 30 minutes at room temperature with goat anti-rabbit peroxidase-conjugate secondary antibody (IgG, H + L; Jackson ImmunoResearch catalog no. 111-035-003) diluted 1:25,000 in 1X PBST with 5% (w/v) non-fat milk. After the incubation, the membranes were again washed with 1X PBST 3 times for 10 minutes each at room temperature, then incubated with ProSignal Pico Spray

chemiluminescence substrate (Prometheus) for 1-2 minutes at room temperature and imaged using a ChemiDoc Imaging System (Bio-Rad). Goat anti-rabbit peroxidase-conjugate secondary antibody (IgG, H + L) was manufacturer validated by antigen-binding assay, western blotting and/or enzyme-linked immunosorbent assay (ELISA). All antibodies against *Xenopus* proteins used in this study were validated by western blotting using *Xenopus* egg extracts.

2.5.15 Preparation of next-generation sequencing libraries

In preparation for next-generation sequencing (NGS), plasmids were replicated as described above. After extraction and ethanol precipitation, samples were resuspended to a concentration of approximately 10-15 ng/µL. 4 µL of recovered DNA were mixed with 200 nM each of forward and reverse primer (see Supplementary Table 1), 200 µM dNTPs, and 2 units of Phusion High-Fidelity DNA Polymerase (NEB) in 1X Phusion HF buffer to a total volume of 100 µL. Amplicons were generated by PCR with initial heating to 98 °C for 30 seconds, followed by 18 cycles of 98 °C for 10 seconds, 55 °C for 30 seconds, and 72 °C for 30 seconds, and a final incubation at 72 °C for 10 minutes. The amplified DNA was purified with a QIAquick PCR Purification Kit (Qiagen) per the manufacturer's instructions and inspected on a 0.8% agarose, 1X TBE gel visualized with SYBR Gold to confirm amplification. Samples were then submitted to Azenta for paired-end next-generation sequencing (Illumina MiSeq, 2x250 bp).

2.5.16 NGS data processing

Raw sequencing reads were cropped to 150 bp using Trimmomatic (v0.39) to avoid later merging failures due to sequencing errors at the ends of amplicons[70]. Reads were then

processed using SIQ (v1.8) with long-read analysis enabled and otherwise default settings, and the SIQPlotteR web tool was used to generate plots[71]. pICL^{clb} contains a single C-C mismatch designed to distinguish reads derived from the two parental strands. These were not deconvoluted in our analysis; instead, the SBS corresponding to the C-C mismatch was manually removed from the SIQ output for pICL^{clb} data sets.

2.5.17 Data availability

Proteomics data were deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD055646 [https://www.ebi.ac.uk/pride/archive/projects/PXD055646]. Next-generation sequencing data were deposited to the NCBI Sequence Read Archive under the BioProject accession number PRJNA1291230 [https://www.ncbi.nlm.nih.gov/sra/PRJNA1291230]. Source data are provided with this paper.

2.5.18 Code availability

Custom code written for this work has been deposited to the California Institute of Technology Research Data Repository (CaltechDATA) with the identifier doi:10.22002/nam70-cvf64[72].

2.6 Acknowledgements

We thank J. Campbell and members of the Semlow, Balskus, Herzon, and Duxin labs for comments on the manuscript. We thank P. Knipscheer for advice on rFANCI-D2 purification and the Tijsterman lab for assistance with implementation of the SIQ software package for NGS data analysis. REV1 and FANCD2 antibodies were gifts from J. Walter

and P. Knipscheer, respectively. FLAG-xlFANCI, Strep-xlFANCD2, and Strep-xlFANCD2^{K562R} expression plasmids were also a gift from J. Walter. D.R.S. is supported by NIH grant no. R01GM151410 and a Shurl and Kay Curci Foundation Research Grant and is a Ronald and JoAnne Willens Scholar. S.B.H. is supported by NIH grant no. R01CA215553. E.P.B. is supported by NIH grant no. R01CA208834 and is a Howard Hughes Medical Institute Investigator. J.P.D. is supported by Novo Nordisk Foundation Grants no. NNF14CC0001 and no. NNF220C0074140. J.W.H.W. was supported by the A*STAR NSS (PhD) predoctoral fellowship.

2.7 Author contributions

M.V.A. performed the experiments described in Figs. 2.1 and 2.3a,b,e,f and Supplementary Figs. 2.1c,d,f,g,j,k, 2.3a-d,g-i, and 2.6d-f. D.R.S. performed the experiments described in Figs. 2.3c,d, 2.4, and 2.5 and Supplementary Figs. 2.1b,e,i, 2.3e,f, 2.4, and 2.5. X.H. performed the experiments described in Figs. 2.6 and 2.7 and Supplementary Figs. 2.6a-c and 2.7. M.V.A., D.R.S., and X.H. performed the experiments described in Fig. 2.2 and Supplementary Fig. 2.2. T.Y.W. and T.F.C. assisted with the mass spectrometry experiments described in Fig. 2.2 and Supplementary Fig. 2.2. E.J.M. performed the data analysis for the mass spectrometry experiments described in Fig. 2.2 and Supplementary Fig. 2.2. K.M.W., Y.A., and S.B.H. supplied the c742 analog. J.W.H.W. and E.P.B provided the DNA duplex containing clb-ICL. S.S.B. and J.P.D provided Pol η and Pol κ antibodies and the pCMV-Sport6-POLH and pET28-6xHis-Polk plasmids. M.V.A and D.R.S. designed the experiments. D.R.S. wrote the manuscript with input from all authors.

2.8 Declaration of interests

E.P.B. is an inventor on the following patents related to detection and inhibition of colibactin biosynthesis: U.S. Patent 11,617,759, U.S. Patent 12,115,176, and U.S. Patent 11,040,951. All other authors have no competing interests to declare.

2.9 References

1 Cao, Y. *et al.* Commensal microbiota from patients with inflammatory bowel disease produce genotoxic metabolites. *Science* **378**, eabm3233 (2022). https://doi.org/10.1126/science.abm3233

2 Wernke, K. M. *et al.* Structure and bioactivity of colibactin. *Bioorg Med Chem Lett* **30**, 127280 (2020). https://doi.org/10.1016/j.bmcl.2020.127280

3 Arthur, J. C. *et al.* Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* **338**, 120-123 (2012). https://doi.org/10.1126/science.1224820

4 Lee-Six, H. *et al.* The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* **574**, 532-537 (2019). https://doi.org/10.1038/s41586-019-1672-7

5 Dziubańska-Kusibab, P. J. *et al.* Colibactin DNA-damage signature indicates mutational impact in colorectal cancer. *Nat Med* **26**, 1063-1069 (2020). https://doi.org/10.1038/s41591-020-0908-2

6 Pleguezuelos-Manzano, C. *et al.* Mutational signature in colorectal cancer caused by genotoxic pks. *Nature* **580**, 269-273 (2020). https://doi.org/10.1038/s41586-020-2080-8

7 Wilson, M. R. *et al.* The human gut bacterial genotoxin colibactin alkylates DNA. *Science* **363** (2019). https://doi.org/10.1126/science.aar7785

8 Xue, M. *et al.* Structure elucidation of colibactin and its DNA cross-links. *Science* **365** (2019). https://doi.org/10.1126/science.aax2685

9 Jiang, Y. *et al.* Reactivity of an Unusual Amidase May Explain Colibactin's DNA Cross-Linking Activity. *J Am Chem Soc* **141**, 11489-11496 (2019). https://doi.org/10.1021/jacs.9b02453

10 Carlson, E. S. *et al.* The specificity and structure of DNA crosslinking by the gut bacterial genotoxin colibactin. *bioRxiv* (2025). https://doi.org/10.1101/2025.05.26.655968

11 Xue, M., Wernke, K. M. & Herzon, S. B. Depurination of Colibactin-Derived Interstrand Cross-Links. *Biochemistry* **59**, 892-900 (2020). https://doi.org/10.1021/acs.biochem.9b01070

12 Vizcaino, M. I. & Crawford, J. M. The colibactin warhead crosslinks DNA. *Nat Chem* **7**, 411-417 (2015). https://doi.org/10.1038/nchem.2221

13 Semlow, D. R. & Walter, J. C. Mechanisms of Vertebrate DNA Interstrand Cross-Link Repair. *Annu Rev Biochem* **90**, 107-135 (2021). https://doi.org/10.1146/annurev-biochem-080320-112510

14 Hodskinson, M. R. *et al.* Alcohol-derived DNA crosslinks are repaired by two distinct mechanisms. *Nature* **579**, 603-608 (2020). https://doi.org/10.1038/s41586-020-2059-5

15 Semlow, D. R., Zhang, J., Budzowska, M., Drohat, A. C. & Walter, J. C. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase. *Cell* **167**, 498-511.e414 (2016). https://doi.org/10.1016/j.cell.2016.09.008

16 Zhang, J. *et al.* DNA interstrand cross-link repair requires replication-fork convergence. *Nat Struct Mol Biol* **22**, 242-247 (2015). https://doi.org/10.1038/nsmb.2956

17 Wu, R. A. *et al.* TRAIP is a master regulator of DNA interstrand crosslink repair. *Nature* **567**, 267-272 (2019). https://doi.org/10.1038/s41586-019-1002-0

18 Fullbright, G., Rycenga, H. B., Gruber, J. D. & Long, D. T. p97 Promotes a Conserved Mechanism of Helicase Unloading during DNA Cross-Link Repair. *Mol Cell Biol* **36**, 2983-2994 (2016). https://doi.org/10.1128/MCB.00434-16

19 Amunugama, R. *et al.* Replication Fork Reversal during DNA Interstrand Crosslink Repair Requires CMG Unloading. *Cell Rep* **23**, 3419-3428 (2018). https://doi.org/10.1016/j.celrep.2018.05.061

20 Garcia-Higuera, I. *et al.* Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol Cell* **7**, 249-262 (2001). https://doi.org/10.1016/s1097-2765(01)00173-3

21 Alcón, P. *et al.* FANCD2-FANCI is a clamp stabilized on DNA by monoubiquitination of FANCD2 during DNA repair. *Nat Struct Mol Biol* **27**, 240-248 (2020). https://doi.org/10.1038/s41594-020-0380-1

22 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). https://doi.org/10.1016/j.molcel.2014.03.015

23 Budzowska, M., Graham, T. G., Sobeck, A., Waga, S. & Walter, J. C. Regulation of the Rev1-pol ζ complex during bypass of a DNA interstrand cross-link. *EMBO J* **34**, 1971-1985 (2015). https://doi.org/10.15252/embj.201490878

24 Long, D. T., Räschle, M., Joukov, V. & Walter, J. C. Mechanism of RAD51-dependent DNA interstrand cross-link repair. *Science* **333**, 84-87 (2011). https://doi.org/10.1126/science.1204258

25 Langevin, F., Crossan, G. P., Rosado, I. V., Arends, M. J. & Patel, K. J. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* **475**, 53-58 (2011). https://doi.org/10.1038/nature10192

26 Rosado, I. V., Langevin, F., Crossan, G. P., Takata, M. & Patel, K. J. Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. *Nat Struct Mol Biol* **18**, 1432-1434 (2011). https://doi.org/10.1038/nsmb.2173

27 García-Rubio, M. L. *et al.* The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. *PLoS Genet* **11**, e1005674 (2015). https://doi.org/10.1371/journal.pgen.1005674

28 Schlacher, K., Wu, H. & Jasin, M. A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. *Cancer Cell* **22**, 106-116 (2012). https://doi.org/10.1016/j.ccr.2012.05.015

29 Luzwick, J. W. *et al.* MRE11-dependent instability in mitochondrial DNA fork protection activates a cGAS immune signaling pathway. *Sci Adv* **7**, eabf9441 (2021). https://doi.org/10.1126/sciadv.abf9441

30 Bona, N. & Crossan, G. P. Fanconi anemia DNA crosslink repair factors protect against LINE-1 retrotransposition during mouse development. *Nat Struct Mol Biol* **30**, 1434-1445 (2023). https://doi.org/10.1038/s41594-023-01067-8

31 van de Kooij, B. *et al.* The Fanconi anemia core complex promotes CtIP-dependent end-resection to drive homologous recombination at DNA double-strand breaks. *bioRxiv* (2023). https://doi.org/10.1101/2023.09.05.556391

32 Bossuet-Greif, N. *et al.* The Colibactin Genotoxin Generates DNA Interstrand Cross-Links in Infected Cells. *mBio* **9** (2018). https://doi.org/10.1128/mBio.02393-17

33 Wernke, K. M. *et al.* Probing Microbiome Genotoxicity: A Stable Colibactin Provides Insight into Structure-Activity Relationships and Facilitates Mechanism of Action Studies. *J Am Chem Soc* **143**, 15824-15833 (2021). https://doi.org/10.1021/jacs.1c07559

34 Liao, S., Matsumoto, Y. & Yan, H. Biochemical reconstitution of abasic DNA lesion replication in Xenopus extracts. *Nucleic Acids Res* **35**, 5422-5429 (2007). https://doi.org/10.1093/nar/gkm552

35 Montagnoli, A. *et al.* A Cdc7 kinase inhibitor restricts initiation of DNA replication and has antitumor activity. *Nat Chem Biol* **4**, 357-365 (2008). https://doi.org/10.1038/nchembio.90

36 Räschle, M. *et al.* Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* **134**, 969-980 (2008). https://doi.org/10.1016/j.cell.2008.08.030

37 Knipscheer, P. *et al.* The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* **326**, 1698-1701 (2009). https://doi.org/10.1126/science.1182372

38 Magnaghi, P. *et al.* Covalent and allosteric inhibitors of the ATPase VCP/p97 induce cancer cell death. *Nat Chem Biol* **9**, 548-556 (2013). https://doi.org/10.1038/nchembio.1313

39 Räschle, M. *et al.* DNA repair. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. *Science* **348**, 1253671 (2015). https://doi.org/10.1126/science.1253671

40 Larsen, N. B. *et al.* Replication-Coupled DNA-Protein Crosslink Repair by SPRTN and the Proteasome in Xenopus Egg Extracts. *Mol Cell* **73**, 574-588.e577 (2019). https://doi.org/10.1016/j.molcel.2018.11.024

41 Wohlschlegel, J. A. *et al.* Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science* **290**, 2309-2312 (2000). https://doi.org/10.1126/science.290.5500.2309

42 Sellés-Baiget, S. *et al.* Catalytic and noncatalytic functions of DNA polymerase κ in translesion DNA synthesis. *Nature Structural & Molecular Biology* (2024). https://doi.org/10.1038/s41594-024-01395-3

43 Kim, H., Yang, K., Dejsuphong, D. & D'Andrea, A. D. Regulation of Rev1 by the Fanconi anemia core complex. *Nat Struct Mol Biol* **19**, 164-170 (2012). https://doi.org/10.1038/nsmb.2222

44 Roy, U. & Schärer, O. D. Involvement of translesion synthesis DNA polymerases in DNA interstrand crosslink repair. *DNA Repair (Amst)* **44**, 33-41 (2016). https://doi.org/10.1016/j.dnarep.2016.05.004

45 Williams, H. L., Gottesman, M. E. & Gautier, J. Replication-independent repair of DNA interstrand crosslinks. *Mol Cell* **47**, 140-147 (2012). https://doi.org/10.1016/j.molcel.2012.05.001

46 Alter, B. P. Cancer in Fanconi anemia, 1927-2001. *Cancer* **97**, 425-440 (2003). https://doi.org/10.1002/cncr.11046

47 Boot, A. *et al.* Characterization of colibactin-associated mutational signature in an Asian oral squamous cell carcinoma and in other mucosal tumor types. *Genome Res* **30**, 803-813 (2020). https://doi.org/10.1101/gr.255620.119

48 Webster, A. L. H. *et al.* Genomic signature of Fanconi anaemia DNA repair pathway deficiency in cancer. *Nature* **612**, 495-502 (2022). https://doi.org/10.1038/s41586-022-05253-4

49 Olivieri, M. *et al.* A Genetic Map of the Response to DNA Damage in Human Cells. *Cell* **182**, 481-496.e421 (2020). https://doi.org/10.1016/j.cell.2020.05.040

50 Stern, H. R., Sefcikova, J., Chaparro, V. E. & Beuning, P. J. Mammalian DNA Polymerase Kappa Activity and Specificity. *Molecules* **24** (2019). https://doi.org/10.3390/molecules24152805

51 Jha, V., Bian, C., Xing, G. & Ling, H. Structure and mechanism of error-free replication past the major benzo[a]pyrene adduct by human DNA polymerase κ. *Nucleic Acids Res* **44**, 4957-4967 (2016). https://doi.org/10.1093/nar/gkw204

52 Gallina, I. *et al.* The ubiquitin ligase RFWD3 is required for translesion DNA synthesis. *Mol Cell* **81**, 442-458.e449 (2021). https://doi.org/10.1016/j.molcel.2020.11.029

53 Shachar, S. *et al.* Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. *EMBO J* **28**, 383-393 (2009). https://doi.org/10.1038/emboj.2008.281

54 Johnson, R. E., Haracska, L., Prakash, S. & Prakash, L. Role of DNA polymerase eta in the bypass of a (6-4) TT photoproduct. *Mol Cell Biol* **21**, 3558-3563 (2001). https://doi.org/10.1128/MCB.21.10.3558-3563.2001

55 Roy, U., Mukherjee, S., Sharma, A., Frank, E. G. & Schärer, O. D. The structure and duplex context of DNA interstrand crosslinks affects the activity of DNA polymerase η. *Nucleic Acids Res* **44**, 7281-7291 (2016). https://doi.org/10.1093/nar/gkw485

56 Kato, N. *et al.* Sensing and Processing of DNA Interstrand Crosslinks by the Mismatch Repair Pathway. *Cell Rep* **21**, 1375-1385 (2017). https://doi.org/10.1016/j.celrep.2017.10.032

57 Yang, W. & Gao, Y. Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. *Annu Rev Biochem* **87**, 239-261 (2018). https://doi.org/10.1146/annurev-biochem-062917-012405

58 Ramsden, D. A., Carvajal-Garcia, J. & Gupta, G. P. Mechanism, cellular functions and cancer roles of polymerase-theta-mediated DNA end joining. *Nat Rev Mol Cell Biol* **23**, 125-140 (2022). https://doi.org/10.1038/s41580-021-00405-2

59 Belan, O. *et al.* POLQ seals post-replicative ssDNA gaps to maintain genome stability in BRCA-deficient cancer cells. *Mol Cell* **82**, 4664-4680.e4669 (2022). https://doi.org/10.1016/j.molcel.2022.11.008

60 Sparks, J. & Walter, J. C. Extracts for Analysis of DNA Replication in a Nucleus-Free System. *Cold Spring Harb Protoc* **2019** (2019). https://doi.org/10.1101/pdb.prot097154

61 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). https://doi.org/10.1007/978-1-61779-998-3_15

62 Lebofsky, R., Takahashi, T., and Walter, J.C. in *Methods in Molecular Biology, DNA Replication* Vol. 521 (ed Sonya Vengrova and Jacob Z. Dalgaard) Ch. 13, 229-252 (Humana Press, 2009).

63 Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372 (2008). https://doi.org/10.1038/nbt.1511

64 Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-2526 (2014). https://doi.org/10.1074/mcp.M113.031591

65 Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319 (2016). https://doi.org/10.1038/nprot.2016.136

66 Deutsch, E. W. *et al.* The ProteomeXchange consortium in 2020: enabling 'big data' approaches in proteomics. *Nucleic Acids Research* (2019). https://doi.org/10.1093/nar/gkz984

67 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740 (2016). https://doi.org/10.1038/nmeth.3901

68 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015). https://doi.org/10.1093/nar/gkv007

69 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **57**, 289-300 (1995). https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

70 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014). https://doi.org/10.1093/bioinformatics/btu170

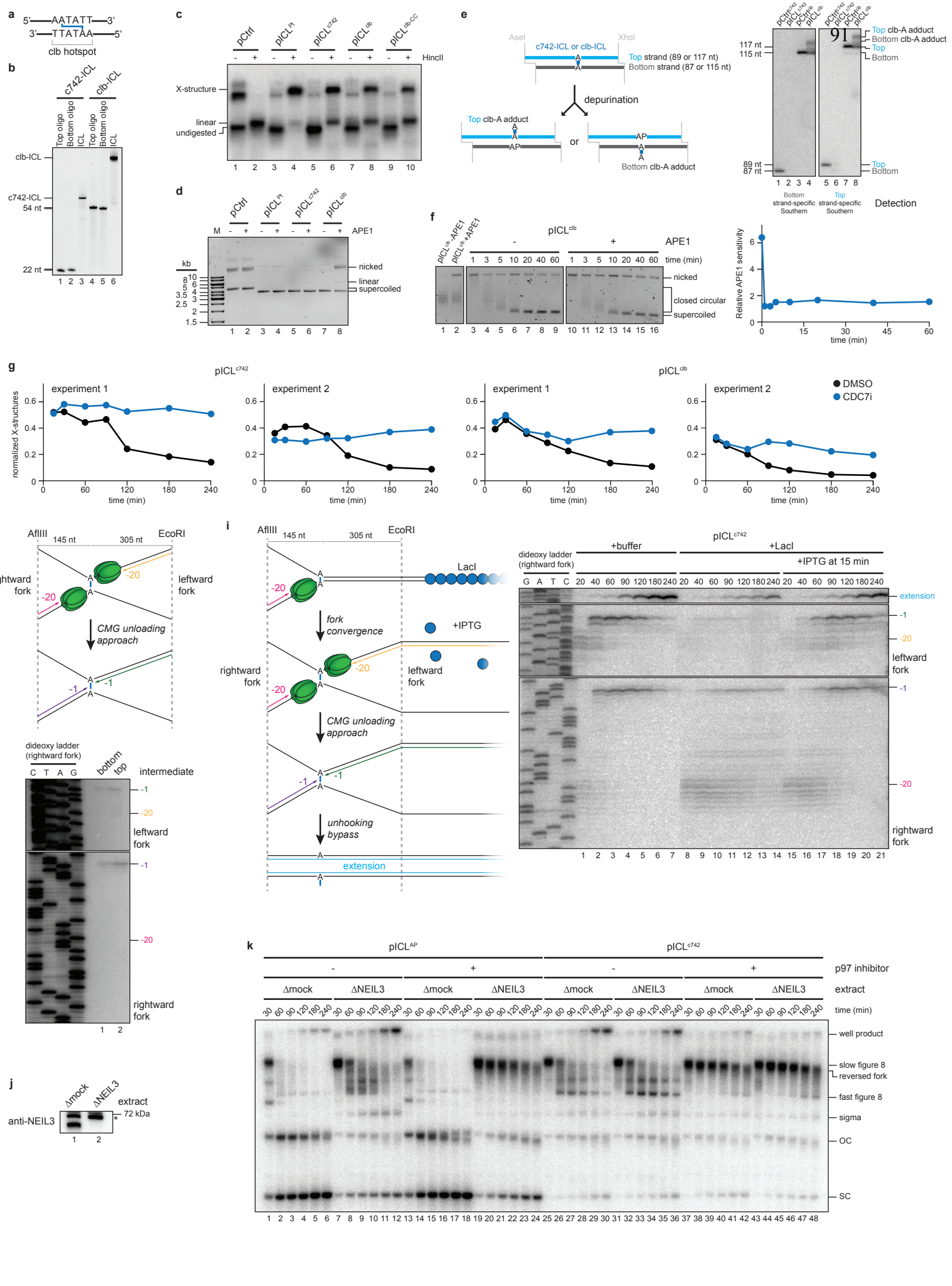
71 van Schendel, R., Schimmel, J. & Tijsterman, M. SIQ: easy quantitative measurement of mutation profiles in sequencing data. *NAR Genom Bioinform* **4**, lqac063 (2022). https://doi.org/10.1093/nargab/lqac063

72 Code for "The Fanconi anemia pathway repairs colibactin-induced DNA interstrand cross-links" (CaltechDATA, 2024).

2.10 Supplementary Information

Name	Sequence	Source
RA302	CTC TCC TGA CTA CTC CCA GTC A	IDT
RA303	GGC GGG ACT ATG GTT GCT	IDT
c742-ICL top	5'-P CCC TGC CCG CAA TAT TGC CGC G	IDT
c742-ICL bottom	5'-P GCA CCG CGG CAA TAT TGC GGG C	IDT
clb-ICL top	5'-P CCC TTC GAT CCC GCG AAT ATT GCG CGA CTC ACT GCA GGG GAA TTG TGA GC	IDT
clb-ICL bottom	5'-P GCA CGC TCA CAA TTC CCC TGC AGT GAG TCG CGC AAT ATT CGC GGG ATC GA	IDT
Pt-ICL top	5’-P CCC TCT TCC GCT CTT CTT TC	IDT
Pt-ICL bottom	5’-P GCA CGA AAG AAG AGC GGA AG	IDT
AP-ICL top	5’-P CCC TCT TCC GCT CdUT CTT TC	IDT
AP-ICL bottom	5’-P GCA CGA AAG AAG AGC GGA AG	IDT
seq ladder primer	CAT GTT TTA CTA GCC AGA TTT TTC CTC CTC TCC TG	IDT

Supplementary Table 2.1 | Oligonucleotide sequences used in this study


a
5'—AATATT—3'
3'—TTATAA—5'
clb hotspot
b
c742-ICL
clb-ICL
Top oligo
Bottom oligo
ICL
clb-ICL
c742-ICL
54 nt
22 nt
c
pCtrl
pICL Pt
pICL c742
pICL clb
pICL clb-CC
HincII
X-structure
linear
undigested
d
pCtrl
pICL Pt
pICL c742
pICL clb
APE1
M
kb
nicked
linear
supercoiled
e
AseI
XhoI
c742-ICL or clb-ICL
Top strand (89 or 117 nt)
Bottom strand (87 or 115 nt)
depurination
Top clb-A adduct
or
AP
Bottom clb-A adduct
117 nt
115 nt
89 nt
87 nt
Top clb-A adduct
Bottom clb-A adduct
Top
Bottom
Bottom strand-specific Southern
Top strand-specific Southern
Detection
f
pICL clb
pICL clb +APE1
APE1
time (min)
nicked
closed circular
supercoiled
Relative APE1 sensitivity
time (min)
g
pICL c742
pICL clb
experiment 1
experiment 2
normalized X-structures
time (min)
DMSO
CDC7i
h
AflIII
145 nt
305 nt
EcoRI
rightward fork
leftward fork
CMG unloading approach
dideoxy ladder (rightward fork)
C T A G
bottom
top
intermediate
i
AflIII
145 nt
305 nt
EcoRI
LacI
+IPTG
fork convergence
CMG unloading approach
unhooking bypass
extension
pICL c742
+buffer
+LacI
+IPTG at 15 min
dideoxy ladder (rightward fork)
G A T C
extension
leftward fork
rightward fork
k
pICL AP
pICL c742
p97 inhibitor
ΔmockΔNEIL3
extract
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
j
Δmock
ΔNEIL3
extract
anti-NEIL3
72 kDa

Supplementary Fig. 2.1 | Preparation and replication of plasmids containing colibactin-induced ICLs

a, Schematic of oligonucleotides containing a 5'-AATATT-3' colibactin alkylation hotspot. The inferred position of the cross-link is indicated.

b, Pairs of annealed oligonucleotides (22 or 54 nt in length) were reacted with c742 or native clb. Purified cross-linked oligos were resolved on a denaturing polyacrylamide gel and visualized with SYBR Gold stain.

c, Plasmids containing the indicated ICLs were linearized with HincII, resolved on a denaturing (alkaline) agarose gel, and detected by Southern blotting, as in Fig. 1c. Note that this method may underestimate the extent of $pICL^{c742}$ and $pICL^{clb}$ cross-linking since it requires extensive incubations, workups, and gel run times, during which the plasmids may undergo depurination.

d, Plasmids containing the indicated ICLs were digested with APE1, resolved on a native agarose gel supplemented with ethidium bromide, and visualized with a UV transilluminator. Note that $pICL^{clb}$ (but not pCtrl, $pICL^{Pt}$, or $pICL^{c742}$) is sensitive to APE1 digestion, indicating the presence of AP sites formed by depurination of the clb-ICL.

e, Left, schematic of fragments produced by AseI and XhoI digestion of plasmids containing intact or depurinated colibactin-induced cross-links. Right, the indicated plasmids were digested with AseI and XhoI, resolved on a denaturing polyacrylamide sequencing gel, and visualized by Southern blotting with 5' end radiolabeled strand-specific probes and autoradiography. Digestion of uncross-linked control plasmids generates only unmodified top and bottom strands. Digestion of $pICL^{clb}$ plasmids generates a mixture of unmodified and adducted top and bottom strands (since depurination can occur at either cross-linked base). Fragments containing intact cross-links migrate slowly and are not detected in portion of the gel used for blotting. Note that neither unmodified nor adducted strands are detected following digestion of $pICL^{c742}$, suggesting that the $pICL^{c742}$ cross-link is largely intact.

f, Left, $pICL^{clb}$ was incubated with pre-mixed HSS and NPE (which does not support replication). The plasmid was recovered at the times indicated, digested with APE1, resolved on a native agarose gel, and visualized by SYBR Gold staining. A major fraction of $pICL^{clb}$ is nicked upon digestion with APE1 (lanes 1 and 2), indicating the presence of AP sites. Following incubation in extract, the increase in nicked species present upon APE1 digestion and undigested samples is dramatically reduced, indicating that AP sites have been repaired. Right, the fraction of plasmid corresponding to open circular, nicked species (OC) in each lane

of the top panel was quantified as the OC intensity relative to the combined intensities of closed circular, supercoiled (SC) and OC species (OC/[OC+SC]). The fold change in the OC fraction upon APE1 digestion is plotted ($[OC/(OC+SC)]_{+APE1}/[OC/(OC+SC)]_{-APE1}$). Note that this APE1-dependent fold change in the OC fraction rapidly declines during incubation in egg extracts, indicating that AP sites are rapidly repaired.

g, Quantification of the Southern blot shown in Fig. 1c (experiment 1) and an independent experimental replicate (experiment 2). The persistence of $pICL^{c742}$ and $pICL^{clb}$ X-structures was quantified as the intensity of X-structures relative to the combined intensities of X-structures and 5.6 kb linear species (X/X+L).

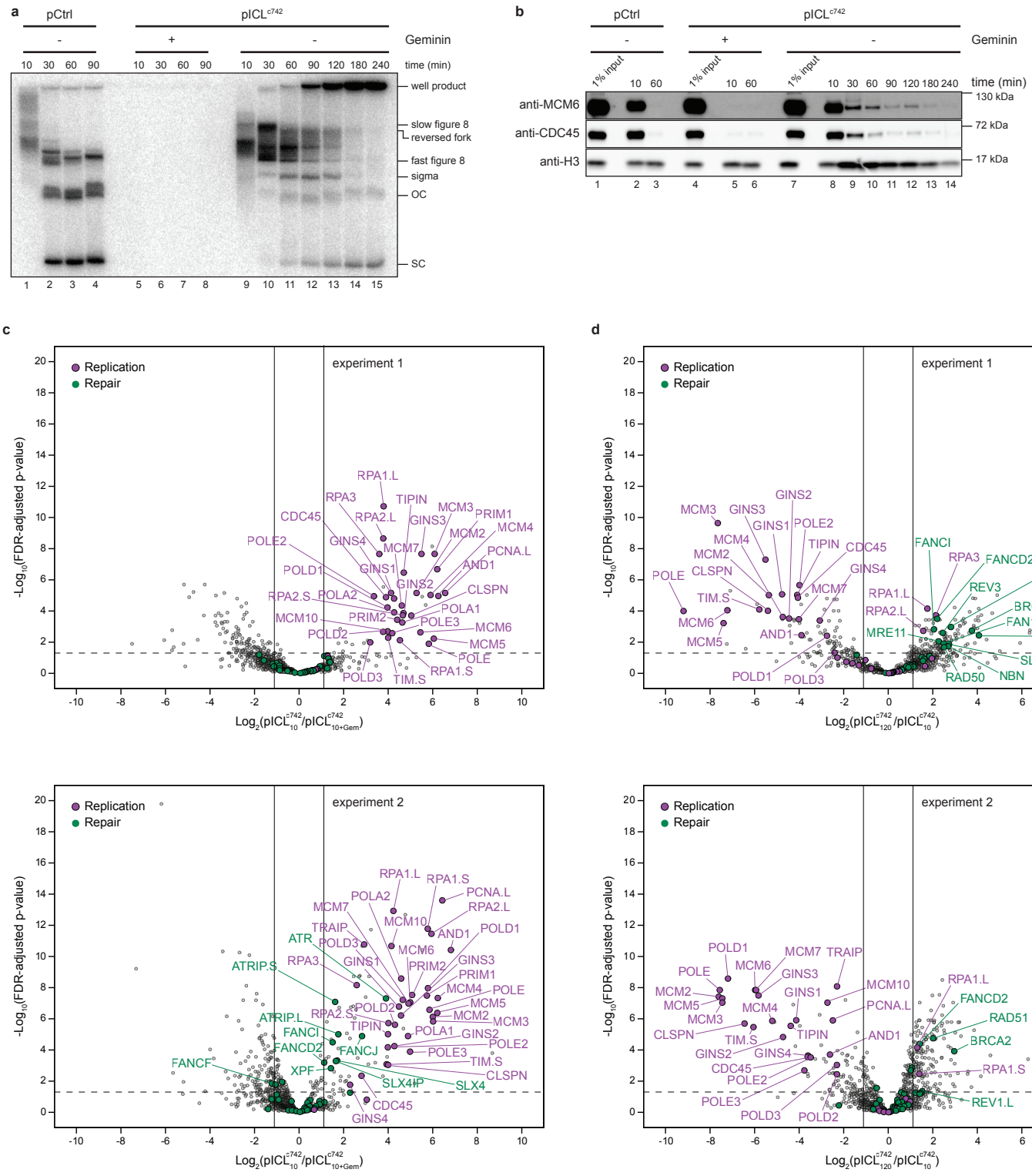
h, Top, schematic of nascent DNA leading strands liberated by AflIII and EcoRI digestion of replication intermediates. Green, CMG helicase. Bottom, $pICL^{c742}$ was replicated in egg extract supplemented with [a-^{32}P]dCTP and, repair intermediates were resolved on a native agarose gel. The slower migrating (top) and faster migrating (bottom) $pICL^{c742}$-specific repair intermediates (i.e. species denoted by red arrowheads in Fig. 1d) were electroeluted from excised gel slices. The intermediates were then digested with AflIII and EcoRI, and the nascent DNA strands were resolved on a denaturing polyacrylamide sequencing gel. Analysis of the nascent DNA indicates that, for both intermediates, the rightward and leftward leading strands have stalled one nucleotide upstream of the cross-link (-1 position).

i, Left, schematic of nascent DNA leading strand intermediates produced in the presence of LacI and IPTG. Binding of LacI (blue circles) to an array of 48 *lacO* sites adjacent to the ICL blocks progression of the leftward replication fork and prevents convergence at the ICL. Addition of IPTG causes LacI to dissociate and enables progression of the leftward fork to the ICL. Nascent lagging strands are omitted for clarity. Right, $pICL^{c742}$ was pre-incubated with buffer or LacI and then replicated in egg extract supplemented with [a-^{32}P]dCTP. Where indicated, IPTG was added 15 min after initiating replication to dissociate LacI. Replication intermediates were digested with AflIII and EcoRI, and nascent DNA leading strands were resolved on a denaturing polyacrylamide sequencing gel. When LacI is bound to $pICL^{c742}$, only the rightward fork encounters the cross-link, and the rightward nascent leading strands stall 20 to 40 nt upstream of the ICL due to the footprint of CMG. Persistent stalling at -20 to -40 is accompanied by a strong reduction in both -1 approach products and full-length extension products, indicating that a single replication fork is insufficient to trigger resolution of the ICL and maturation of the leading strands. Addition of IPTG enables the leftward fork to reach the ICL. Arrival of the leftward fork triggers approach of the nascent leading strands up to the

ICL and subsequent formation of full-length extension products. This result indicates that pICL^{c742} unhooking requires convergence of two replication forks on either side of the ICL.

j, NEIL3 immunodepletion. The extracts used in the replication reactions shown in **k** were blotted for NEIL3. Asterisk indicates non-specific band.

k, pICL^{AP} and pICL^{c742} were replicated in mock- or NEIL3-depleted egg extracts supplemented with [a-^{32}P]dCTP and p97 inhibitor, as indicated. Replication intermediates were analyzed as in Fig. 1d. SC, supercoiled; OC, open circular. Whereas depletion of NEIL3 dramatically inhibits repair of pICL^{AP}, it has no discernible impact on repair of pICL^{c742}. This result indicates that NEIL3 does not contribute to the resolution of colibactin-induced ICLs.


a
pCtrl
pICL c742
Geminin
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
b
pCtrl
pICL c742
Geminin
1% input
time (min)
anti-MCM6
anti-CDC45
anti-H3
130 kDa
72 kDa
17 kDa
c
d
Replication
Repair
experiment 1
experiment 2
-Log10(FDR-adjusted p-value)
Log2(pICL c742 10/pICL c742 10+Gem)
Log2(pICL c742 120/pICL c742 10)

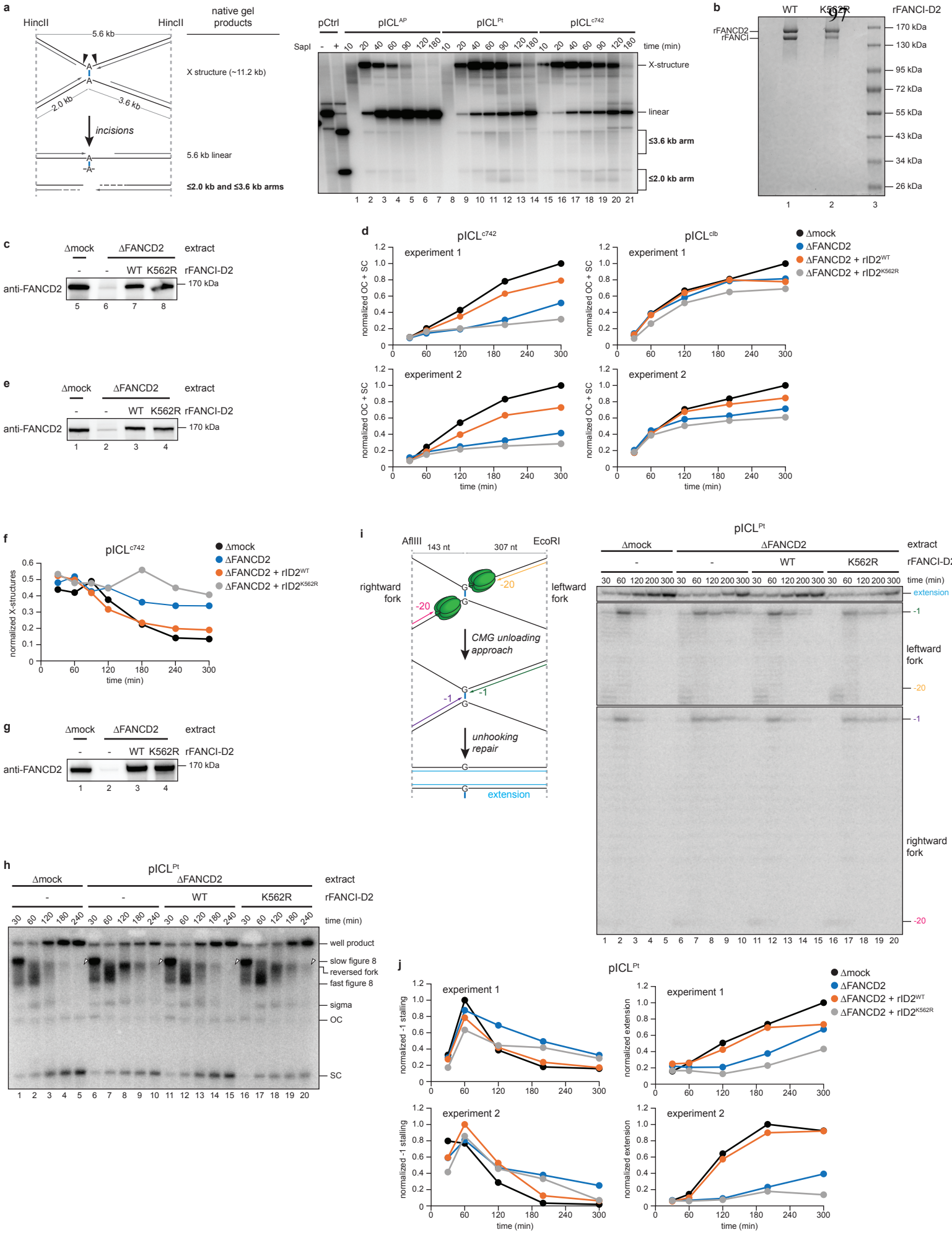
Supplementary Fig. 2.2 | Mass spectrometry analyses of c742-ICL repair

a, Aliquots of the replication reactions described in Fig. 2, experiment 1, were supplemented with [a-^{32}P]dCTP, and replication intermediates were analyzed as in Fig. 1d. SC, supercoiled; OC, open circular.

b, Samples of chromatin-associated proteins recovered from the replication reactions described in Fig. 2, experiment 1, were resolved by SDS-PAGE, and the indicated proteins were detected by immunoblotting.

c, Volcano plots comparing the recruitment of proteins to pICL^{c742} in replicating extract (pICL) vs non-replicating (pICL$_{+Gem}$) extract at 10 minutes following replication initiation (top, experiment 1; bottom, experiment 2). Estimated fold changes and FDR-adjusted p-values were calculated as in Fig. 2b. Labeled proteins of interest meet significance criteria ($|Log_2FC| \geq 1$ and $p \leq 0.05$).

d, Volcano plots comparing the recruitment of proteins to pICL^{c742} at 120 minutes vs at 10 minutes following replication initiation (top, experiment 1; bottom, experiment 2). Estimated fold changes and FDR-adjusted p-values were calculated as in Fig. 2b. Labeled proteins of interest meet significance criteria ($|Log_2FC| \geq 1$ and $p \leq 0.05$).


a
HincII
HincII
5.6 kb
native gel products
X structure (~11.2 kb)
2.0 kb
3.6 kb
incisions
5.6 kb linear
≤2.0 kb and ≤3.6 kb arms
pCtrl
pICL^AP
pICL^Pt
pICL^c742
SapI
time (min)
X-structure
linear
≤3.6 kb arm
≤2.0 kb arm
b
WT
K562R
rFANCI-D2
rFANCD2
rFANCI
170 kDa
130 kDa
95 kDa
72 kDa
55 kDa
43 kDa
34 kDa
26 kDa
c
ΔmockΔFANCD2extract
WT K562R rFANCI-D2
anti-FANCD2
170 kDa
d
pICL^c742
pICL^clb
experiment 1
experiment 2
normalized OC + SC
time (min)
Δmock
ΔFANCD2
ΔFANCD2 + rID2^WT
ΔFANCD2 + rID2^K562R
e
ΔmockΔFANCD2extract
WT K562R rFANCI-D2
anti-FANCD2
170 kDa
f
pICL^c742
normalized X-structures
time (min)
Δmock
ΔFANCD2
ΔFANCD2 + rID2^WT
ΔFANCD2 + rID2^K562R
g
ΔmockΔFANCD2extract
WT K562R rFANCI-D2
anti-FANCD2
170 kDa
h
pICL^Pt
ΔmockΔFANCD2extract
WT K562R rFANCI-D2
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
i
AflIII
143 nt
307 nt
EcoRI
rightward fork
leftward fork
-20
-1
CMG unloading approach
unhooking repair
extension
pICL^Pt
ΔmockΔFANCD2extract
WT K562R rFANCI-D2
time (min)
extension
-1
leftward fork
-20
-1
rightward fork
-20
j
pICL^Pt
experiment 1
experiment 2
normalized -1 stalling
normalized extension
time (min)
Δmock
ΔFANCD2
ΔFANCD2 + rID2^WT
ΔFANCD2 + rID2^K562R

Supplementary Fig. 2.3 | Colibactin- and cisplatin-induced ICLs are unhooked by the FA pathway

a, Left, schematic of incision products generated during nucleolytic ICL repair. Digestion with HincII produces X-structures corresponding to cross-linked plasmids and linear species corresponding to unhooked plasmids. Nucleolytic incisions flanking the ICL produce double-strand break fragments that are resolved on a native agarose gel as ~2.0 and ~3.6 kb species. Right, $pICL^{AP}$, $pICL^{Pt}$, and $pICL^{c742}$ were replicated in egg extract supplemented with [a-^{32}P]dCTP. Recovered replication intermediates were digested with HincII, resolved by native agarose gel electrophoresis, and visualized by autoradiography.

b, Purified recombinant *Xenopus laevis* rFANCI-FANCD2 protein complexes were resolved by SDS-PAGE and visualized by staining with Coomassie stain. rFANCI harbors a FLAG tag. rFANCD2 harbors a Strep tag. rFANCI-$FANCD2^{K562R}$ harbors a FANCD2 K562R mutation that blocks monoubiquitylation and inhibits ICL repair in egg extracts[37].

c, FANCD2 immunodepletion. The extracts used in the replication reactions shown in Fig. 3b (right) and Fig. 3e (bottom) were blotted for FANCD2.

d, Quantification of open circular and supercoiled repair products from the experiment shown in Fig. 3b (experiment 1) and an independent experimental replicate (experiment 2). The summed intensities of open circular and supercoiled species (OC+SC) were normalized to the maximum of the summed open circular and supercoiled intensities across all samples.

e, FANCD2 immunodepletion. The extracts used in the replication reactions shown in Fig. 3c were blotted for FANCD2.

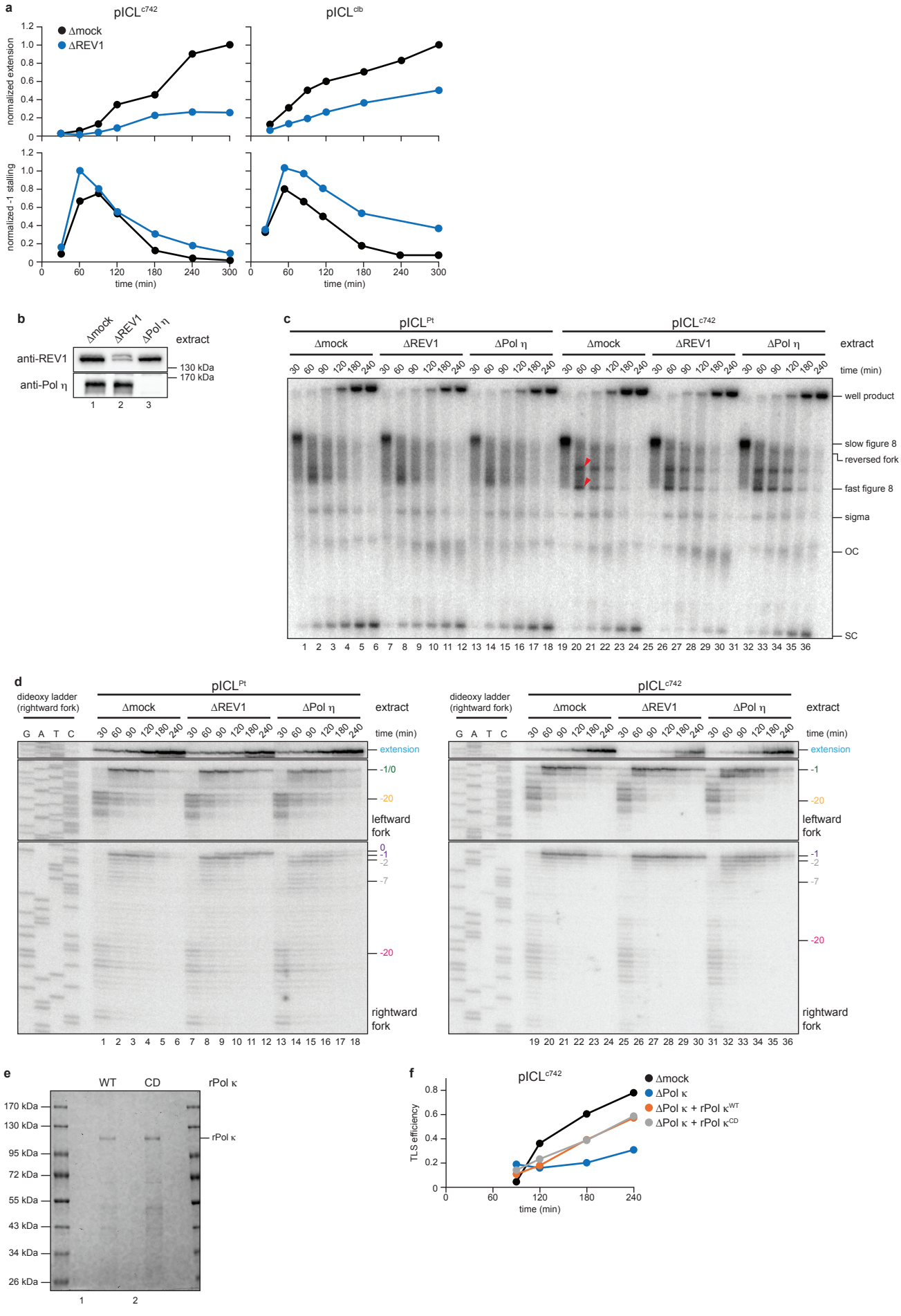
f, Quantification of an independent replicate of the experiment shown in Fig. 3c. The persistence of $pICL^{c742}$ X-structures was quantified as the intensity of X-structures relative to the combined intensities of X-structures and 5.6 kb linear species (X/X+L).

g, FANCD2 immunodepletion. The extracts used in the replication reactions shown in **h** and **i** were blotted for FANCD2.

h, $pICL^{Pt}$ was replicated in mock- or FANCD2-depleted extracts supplemented with [a-^{32}P]dCTP and rFANCI-$FANCD2^{WT}$ or rFANCI-$FANCD2^{K562R}$, as indicated. Replication intermediates were analyzed as in Fig. 1d. SC, supercoiled; OC, open circular. White arrowheads indicate reversed fork intermediates.

i, Nascent DNA strands from the pICL^{Pt} replication reactions shown in **h** were digested with AflIII and EcoRI and analyzed as in Fig. 1f.

j, Quantification of -1 stall and strand extension products shown in **i**. The intensities of leftward fork -1 stall products (left panels) or extension products (right panels) were normalized to maximum -1 stall or extension product intensity, respectively, across all samples.


a
pICL^c742
pICL^clb
Δmock
ΔREV1
normalized extension
normalized -1 stalling
time (min)
b
Δmock
ΔREV1
ΔPol η
extract
anti-REV1
anti-Pol η
130 kDa
170 kDa
c
pICL^Pt
pICL^c742
Δmock
ΔREV1
ΔPol η
extract
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
d
dideoxy ladder (rightward fork)
G A T C
extension
-1/0
-20
leftward fork
0
-1
-2
-7
rightward fork
e
WT
CD
rPol κ
170 kDa
130 kDa
95 kDa
72 kDa
55 kDa
43 kDa
34 kDa
26 kDa
f
TLS efficiency
Δmock
ΔPol κ
ΔPol κ + rPol κ^WT
ΔPol κ + rPol κ^CD

Supplementary Fig. 2.4 | Pol η, REV1, and Pol κ promote translesion synthesis during colibactin-induced ICL repair

a, Quantification of -1 stall (bottom) and strand extension (top) products for an independent replicate of the experiment shown in Fig. 4c. The intensities of rightward fork -1 stall products or extension products were normalized to maximum -1 stall or extension product intensity, respectively, across all samples.

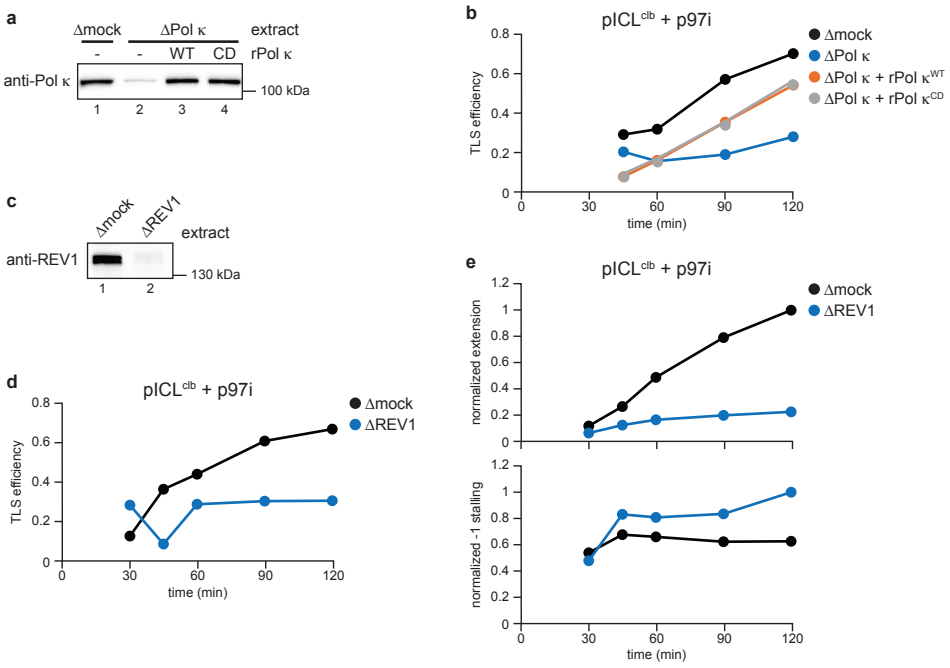
b, REV1 and Pol η immunodepletion. The extracts used in the replication reactions shown in **c** and **d** were blotted for REV1 and Pol η.

c, $pICL^{Pt}$ and $pICL^{c742}$ were replicated in mock-, REV1-, or Pol η-depleted extracts supplemented with [a-^{32}P]dCTP, and replication intermediates were analyzed as in Fig. 1d. SC, supercoiled; OC, open circular. Red arrowheads indicate top and bottom intermediates formed during replication of $pICL^{c742}$.

d, Nascent DNA strands from the $pICL^{Pt}$ and $pICL^{c742}$ replication reactions shown in **c** were digested with AflIII and EcoRI and analyzed as in Fig. 1f.

e, Purified His-tagged recombinant *Xenopus laevis* rPol κ proteins were resolved by SDS-PAGE and visualized by staining with Coomassie stain. rPol $\kappa^{D198A/E199A}$ (referred to as rPol κ^{CD}) harbors D198E and E199A mutations that eliminate polymerase activity[42].

f, TLS efficiencies for an independent replicate of the experiment shown in Fig. 4f were quantified as the intensity of supercoiled plasmid relative to the combined intensities of open circular and supercoiled plasmids (SC/OC+SC).



Supplementary Fig. 2.5 | Pol κ and REV1 promote translesion synthesis past a colibactin-induced monoadduct

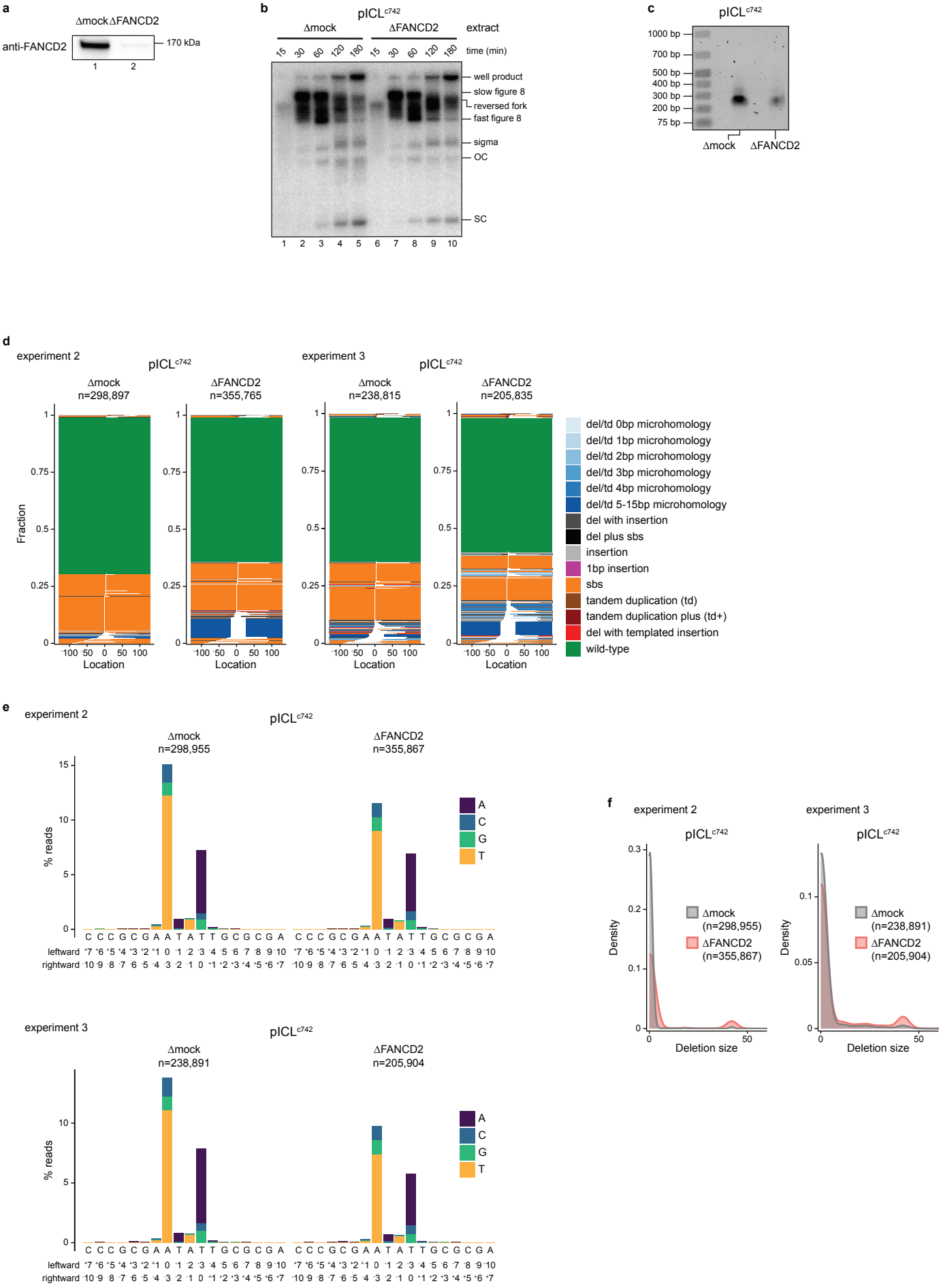
a, Pol κ immunodepletion. The extracts used in the replication reaction shown in Fig. 5a were blotted for Pol κ.

b, TLS efficiencies for an independent replicate of the experiment shown in Fig. 5a were quantified as the intensity of supercoiled plasmid relative to the combined intensities of open circular and supercoiled plasmids (SC/OC+SC).

c, REV1 immunodepletion. The extracts used in the replication reactions shown in Fig. 5c were blotted for REV1.

d, TLS efficiencies for an independent replicate of the experiment shown in Fig. 5c were quantified as in **b**.

e, Quantification of -1 stall and strand extension products for an independent replicate of the experiment shown in Fig. 5e. The intensities of rightward fork -1 stall products (bottom) or extension products (top) were normalized to maximum -1 stall or extension product intensity, respectively, across all samples.


a
ΔmockΔFANCD2
anti-FANCD2
170 kDa
1
2
b
pICL^c742
Δmock
ΔFANCD2
extract
15 30 60 120 180 15 30 60 120 180
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10
c
pICL^c742
1000 bp
700 bp
500 bp
400 bp
300 bp
200 bp
75 bp
Δmock
ΔFANCD2
d
experiment 2
pICL^c742
Δmock
n=298,897
ΔFANCD2
n=355,765
experiment 3
pICL^c742
Δmock
n=238,815
ΔFANCD2
n=205,835
Fraction
Location
del/td 0bp microhomology
del/td 1bp microhomology
del/td 2bp microhomology
del/td 3bp microhomology
del/td 4bp microhomology
del/td 5-15bp microhomology
del with insertion
del plus sbs
insertion
1bp insertion
sbs
tandem duplication (td)
tandem duplication plus (td+)
del with templated insertion
wild-type
e
experiment 2
pICL^c742
Δmock
n=298,955
ΔFANCD2
n=355,867
% reads
A
C
G
T
leftward
rightward
experiment 3
pICL^c742
Δmock
n=238,891
ΔFANCD2
n=205,904
f
experiment 2
pICL^c742
Density
Δmock
(n=298,955)
ΔFANCD2
(n=355,867)
Deletion size
experiment 3
pICL^c742
Density
Δmock
(n=238,891)
ΔFANCD2
(n=205,904)
Deletion size

Supplementary Fig. 2.6 | Independent experimental replicates demonstrating mutagenic repair of colibactin-induced ICLs by the FA pathway

a, FANCD2 immunodepletion. Extracts used in replication reactions described in Fig. 6 were blotted for FANCD2.

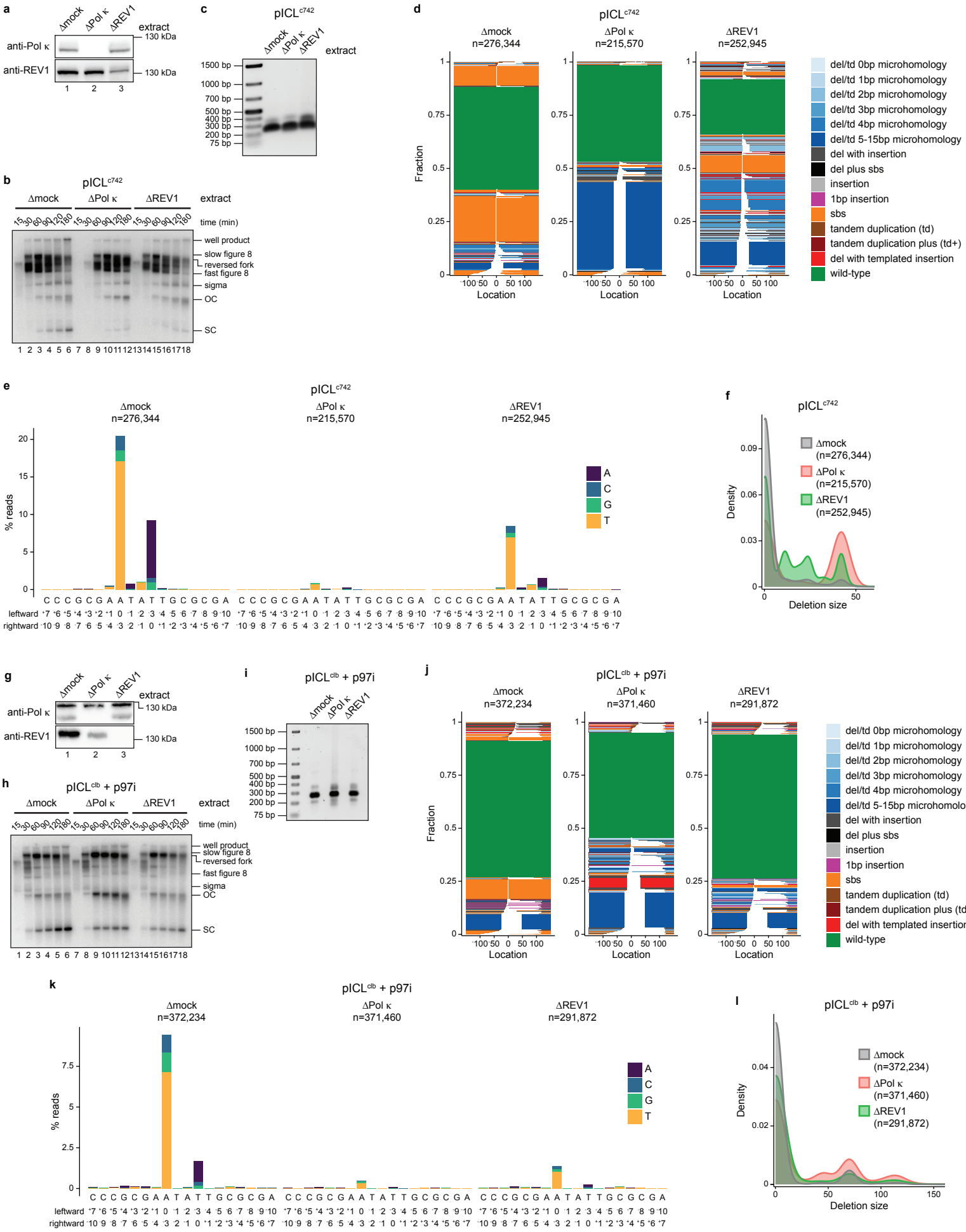
b, pICL^{c742} replication reactions used in the preparation of sequencing libraries described in Fig. 6 were supplemented with [a-^{32}P]dCTP and analyzed as in Fig. 1d. SC, supercoiled; OC, open circular.

c, PCR amplicons generated from the pICL^{c742} replication reactions described in Fig. 6 were resolved on a native agarose gel and visualized with SYBR Gold stain. The expected length of the PCR amplicons is 240 bp.

d, Color-coded repair outcomes are depicted as tornado plots for two additional independent sets of pICL^{c742} replication reactions performed with mock- or FANCD2-depleted extract (experiment 2 and experiment 3). White spaces indicate deletions. n indicates the number of mapped sequencing reads obtained for each condition.

e, The percentage of reads containing single base substitutions at the indicated positions relative to the ICL is plotted for the additional data sets described in **d**. The reference sequence is shown with numbers indicating the positions for the rightward and leftward nascent leading strands relative to the site of cross-linking ("0"). n indicates the number of mapped sequencing reads obtained for each condition.

f, The fraction of deletions spanning a given length is plotted for the additional data sets described in **d**.


a
Δmock
ΔPol κ
ΔREV1
extract
anti-Pol κ
anti-REV1
130 kDa
130 kDa
1 2 3
c
pICL^c742
Δmock
ΔPol κ
ΔREV1
extract
1500 bp
1000 bp
700 bp
500 bp
400 bp
300 bp
200 bp
75 bp
b
pICL^c742
Δmock
ΔPol κ
ΔREV1
extract
time (min)
15 30 60 90 120 180
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
d
pICL^c742
Δmock
n=276,344
ΔPol κ
n=215,570
ΔREV1
n=252,945
Fraction
Location
del/td 0bp microhomology
del/td 1bp microhomology
del/td 2bp microhomology
del/td 3bp microhomology
del/td 4bp microhomology
del/td 5-15bp microhomology
del with insertion
del plus sbs
insertion
1bp insertion
sbs
tandem duplication (td)
tandem duplication plus (td+)
del with templated insertion
wild-type
e
pICL^c742
Δmock
n=276,344
ΔPol κ
n=215,570
ΔREV1
n=252,945
% reads
A
C
G
T
C C C G C G A A T A T T G C G C G A
leftward
rightward
f
pICL^c742
Density
Deletion size
Δmock
(n=276,344)
ΔPol κ
(n=215,570)
ΔREV1
(n=252,945)
g
Δmock
ΔPol κ
ΔREV1
extract
anti-Pol κ
anti-REV1
130 kDa
130 kDa
1 2 3
h
pICL^clb + p97i
Δmock
ΔPol κ
ΔREV1
extract
time (min)
15 30 60 90 120 180
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
i
pICL^clb + p97i
Δmock
ΔPol κ
ΔREV1
1500 bp
1000 bp
700 bp
500 bp
400 bp
300 bp
200 bp
75 bp
j
pICL^clb + p97i
Δmock
n=372,234
ΔPol κ
n=371,460
ΔREV1
n=291,872
Fraction
Location
del/td 0bp microhomology
del/td 1bp microhomology
del/td 2bp microhomology
del/td 3bp microhomology
del/td 4bp microhomology
del/td 5-15bp microhomology
del with insertion
del plus sbs
insertion
1bp insertion
sbs
tandem duplication (td)
tandem duplication plus (td+)
del with templated insertion
wild-type
k
pICL^clb + p97i
Δmock
n=372,234
ΔPol κ
n=371,460
ΔREV1
n=291,872
% reads
A
C
G
T
C C C G C G A A T A T T G C G C G A
leftward
rightward
l
pICL^clb + p97i
Density
Deletion size
Δmock
(n=372,234)
ΔPol κ
(n=371,460)
ΔREV1
(n=291,872)

Supplementary Fig. 2.7 | Next-generation sequencing shows the mutagenicity of TLS past c742- and clb-induced adducts

a, REV1 and Pol κ immunodepletion. Extracts used in replication reactions described in Fig. 7a-c were blotted for Pol κ, and REV1.

b, pICL^{c742} replication reactions used in the preparation of sequencing libraries described in Fig. 7a-c were supplemented with [a-^{32}P]dCTP and analyzed as in Fig. 1d. SC, supercoiled; OC, open circular.

c, PCR amplicons generated from the pICL^{c742} replication reactions described in Fig. 7a-c were resolved on a native agarose gel and visualized with SYBR Gold stain. The expected length of the PCR amplicons is 240 bp.

d, Color-coded repair outcomes are depicted as tornado plots for an additional independent set of pICL^{c742} replication reactions performed with mock-, Pol κ- or REV1-depleted extract. White spaces indicate deletions. n indicates the number of mapped sequencing reads obtained for each condition.

e, The percentage of reads containing single base substitutions at the indicated positions relative to the ICL is plotted for the additional data set described in **d**. The reference sequence is shown with numbers indicating the positions for the rightward and leftward nascent leading strands relative to the site of cross-linking ("0"). n indicates the number of mapped sequencing reads obtained for each condition.

f, The fraction of deletions spanning a given length is plotted for the additional data set described in **d**.

g, REV1 and Pol κ immunodepletion. Extracts used in replication reactions described in Fig. 7e-g were blotted for Pol κ, and REV1.

h, pICL^{clb} replication reactions used in the preparation of sequencing libraries described in Fig. 7e-g were supplemented with [a-^{32}P]dCTP and analyzed as in Fig. 1d. SC, supercoiled; OC, open circular.

i, PCR amplicons generated from the pICL^{clb} replication reactions described in Fig. 7e-g were resolved on a native agarose gel and visualized with SYBR Gold stain. The expected length of the PCR amplicons is 268 bp.

j, Color-coded repair outcomes are depicted as tornado plots for an additional independent set of pICL^{clb} replication reactions performed with mock-, Pol κ- or REV1-depleted extract. White spaces indicate deletions. n indicates the number of mapped sequencing reads obtained for each condition.

k, The percentage of reads containing single base substitutions at the indicated positions relative to the ICL is plotted for the additional data set described in **j**. The reference sequence is shown with numbers indicating the positions for the rightward and leftward nascent leading strands relative to the site of cross-linking (“0”). n indicates the number of mapped sequencing reads obtained for each condition.

l, The fraction of deletions spanning a given length is plotted for the additional data set described in **j**.

Chapter 3

ATM PROMOTES NUCLEOLYTIC PROCESSING OF REVERSED REPLICATION FORKS DURING DNA INTERSTRAND CROSS-LINK REPAIR

Maria Altshuller[1], Victoria A. MacKrell[1], Jasmine Tzeng[1], Richa Nigam[1], Jade K. Konsler[1], Ting-Yu Wang[2,3], Emma J. Grogan[4], Obehi E. Ozua[4], Judith L. Campbell[1,2], James M. Dewar[4], Tsui-Fen Chou[2,3], and Daniel R. Semlow[1,*]

Affiliations

[1]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA

[2]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA

[3]Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, Pasadena, CA, USA

[4]Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN, USA

*Corresponding author (dsemlow@caltech.edu)

This chapter consists of a manuscript that is currently in review. The preprint version is available at:

Altshuller, M. *et al.* ATM promotes reversed fork processing during DNA interstrand cross-link repair. *bioRxiv* (2025). https://doi.org/10.1101/2025.10.08.680157

3.1 Abstract

During replication-coupled DNA interstrand cross-link (ICL) repair, fork reversal is thought to enable the Fanconi anemia (FA) pathway to resolve the ICL through nucleolytic incisions. Subsequent fork restoration then allows extension of nascent DNA past the damaged site. Despite the importance of these remodeling events, their regulation remains poorly understood. Here, we use cell-free *Xenopus* egg extracts and human cells to investigate fork dynamics during ICL repair by the FA pathway. We find that the ataxia-telangiectasia mutated (ATM) kinase is activated concomitantly with fork reversal and promotes resection of reversed fork intermediates. ATM enhances chromatin recruitment of the DNA2-WRN complex, which acts alongside EXO1 to resect reversed forks formed during ICL repair. Moreover, inhibition of protein phosphatase 2A (PP2A) leads to ATM hyperactivation, accelerated reversed fork processing, and accumulation of aberrant end-joining products, implying that PP2A restrains ATM signaling to limit reversed fork resection. Taken together, this work identifies reversed forks as substrates for ATM activation and reveals a phospho-regulatory circuit that controls reversed fork processing.

3.2 Introduction

DNA interstrand cross-links (ICLs) covalently link the two strands of DNA and thereby block strand separation. These lesions are produced by endogenous cellular metabolites (e.g. formaldehyde, crotonaldehyde, abasic sites), microbiome-derived secondary metabolites (e.g. colibactin), and chemotherapeutics (e.g. cisplatin, nitrogen mustard derivatives, psoralens)[1,2]. Mutations in genes associated with ICL repair cause human diseases including the chronic kidney disease karyomegalic interstitial nephritis and the bone marrow failure and cancer predisposition syndrome Fanconi anemia[1,3]. ICLs represent formidable impediments to DNA replication and can lead to deletions, chromosomal rearrangements, and chromosome missegregation. Consequently, even a single unrepaired ICL can be lethal to a eukaryotic cell[4].

Among the multiple mechanisms that cells have evolved to resolve ICLs, the Fanconi anemia (FA) pathway is thought to be the most important for ICL resistance in proliferating cells[1]. The FA pathway is coupled to DNA replication and can be triggered by the convergence and stalling of replisomes on either side of an ICL (Fig. 3.1a)[5]. Fork convergence induces TRAIP-dependent ubiquitylation and p97-dependent unloading of the replicative CDC45-MCM2-7-GINS (CMG) helicase[6-8]. One of the two forks abutting the ICL undergoes extensive reversal, in which reannealing of the parental DNA strands leads to the displacement and annealing of nascent strands and the formation of a four-helix junction[9]. Interestingly, during ICL repair in *Xenopus* egg extracts, this fork reversal is coupled to resection of the nascent lagging strand, resulting in a regressed arm that comprises predominantly single-stranded DNA[9]. Although a requirement for fork reversal during ICL repair has not been proven, reversed forks appear to represent on-pathway ICL

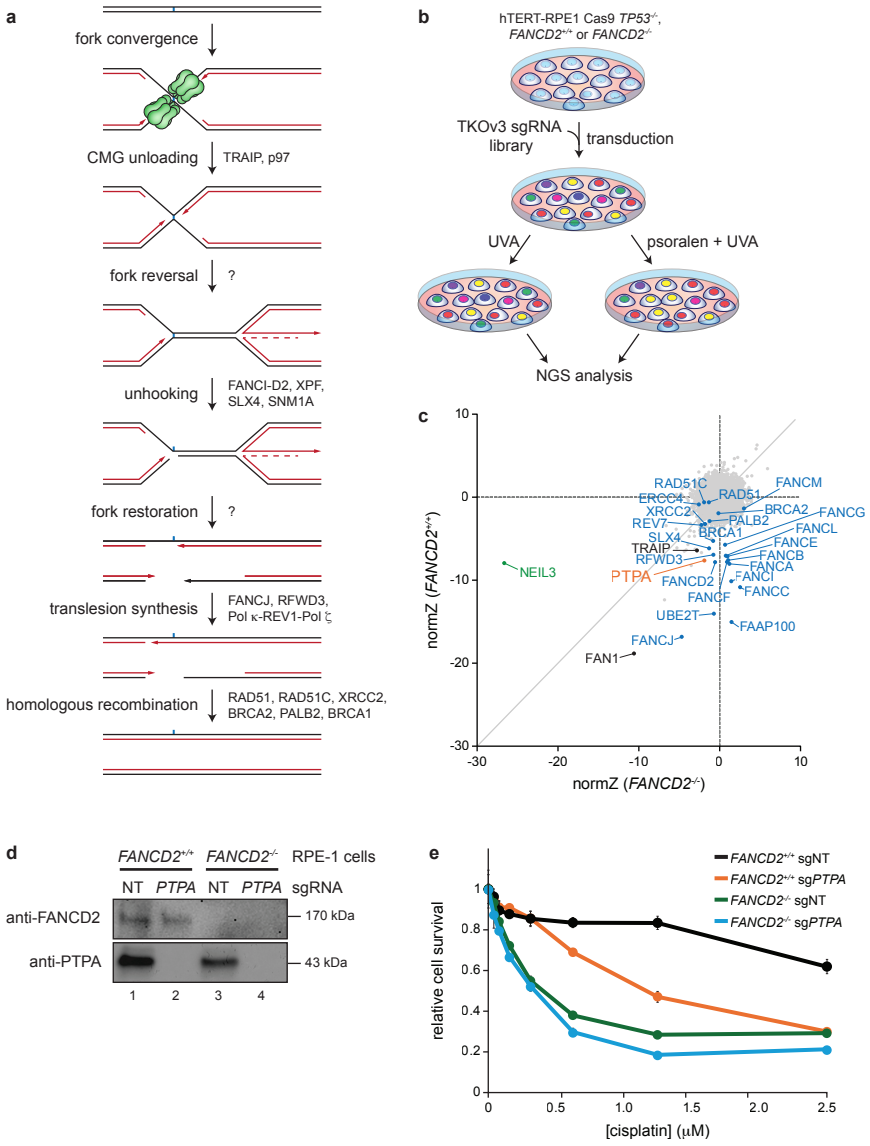


Fig. 3.1 | Genome-wide CRISPR dropout screens implicate PTPA in DNA interstrand cross-link repair

a, Model of replication-coupled ICL repair by the FA pathway.

b, Schematic describing CRISPR dropout screens using human *FANCD2*$^{+/+}$ or *FANCD2*$^{-/-}$ RPE-1 cells.

c, Biplot of normZ scores relating viability in psoralen and UVA treated *FANCD2*$^{+/+}$ (y-axis) and *FANCD2*$^{-/-}$ (x-axis) RPE-1 cells.

d, Immunoblot demonstrating efficiency of PTPA depletion in *FANCD2*$^{+/+}$ and *FANCD2*$^{-/-}$ RPE-1 CRISPR knockout cell pools. Cells were transduced with lentivirus carrying non-targeting (NT) and *PTPA*-targeting sgRNAs.

e, Relative RPE-1 cell proliferation after cisplatin treatment. *FANCD2*$^{+/+}$ and *FANCD2*$^{-/-}$ RPE-1 cells were transduced with sgRNAs as in **d** and treated with cisplatin. Cell proliferation was measured after 5 days using CellTiter-Glo reagent. Data represent the mean ± standard deviation from three technical replicates.

repair intermediates. Additionally, fork reversal establishes the ICL within a duplex at a three-helix junction. This configuration represents the preferred substrate for the XPF-ERCC1 3' flap endonuclease, which incises one parental DNA strand 5' of the ICL[10]. A second nuclease, such as SNM1A, mediates 5' to 3' expansion of the XPF-induced nick past the ICL to release ("unhook") the cross-link and generate a DNA double strand break (DSB). Subsequent steps in ICL repair are less well-defined. However, restoration of the reversed fork likely is required to deliver the 3' end of the nascent leading strand up to the ICL, allowing translesion synthesis past the unhooked ICL adduct ("ICL remnant") and completion of replication on one sister chromatid[11,12]. The two-ended DSB produced upon unhooking and fork restoration is then repaired by homologous recombination using the fully-replicated sister chromatid as a template[13].

While current evidence indicates that replication fork reversal and restoration are integral steps in ICL repair by the FA pathway, little is known about the enzymes that execute these fork dynamics or how they are regulated. The enzyme that mediates fork reversal in the FA pathway has not been identified, though the FANCM translocase is a likely candidate[14]. The factors responsible for fork restoration during ICL repair are also unknown, but this process may depend on one or more RECQ family translocases. The RECQ family members WRN and BLM can perform restoration of model reversed fork substrates *in vitro*[15,16], and RECQ1 promotes fork restoration in response to camptothecin-induced replication stress[17]. Fork restoration during ICL repair may also depend on nuclease

activities. For instance, the DNA2 nuclease-helicase and WRN helicase are known to act cooperatively to perform 5' to 3' resection of reversed fork regressed arms and promote replication restart following replication stress[18]. Currently, the mechanisms that regulate fork reversal and restoration are not fully understood.

Emerging evidence suggests that the ataxia-telangiectasia mutated (ATM) kinase regulates the replication stress response[19-21]. ATM is activated at DSB ends through interactions with the MRE11-RAD50-NBS1 (MRN) nuclease complex and promotes resection and RAD51 filament formation at DSBs[22-24]. ATM thereby stimulates homology-directed repair of two-ended DSBs as well as one-ended DSBs generated upon replication fork collapse. Intriguingly, genetic interactions between ATM and fork reversal factors provide indirect evidence that reversed fork structures may also activate ATM[25,26]. However, these studies induced fork reversal using agents that also produce DSBs and/or oxidative stress, the latter of which can activate ATM independently of DSBs[27,28]. Direct evidence that fork reversal triggers ATM activation is therefore lacking. The functional relevance of this potential ATM activation at reversed forks is also unclear, although a recent report indicates that ATM negatively regulates MRN to maintain fork protection and enable recovery form replication stress[25]. Whether, analogous to its function at canonical DSBs, ATM also promotes resection of reversed forks is unknown. Additionally, how ATM is itself regulated in response to replication stress has not been determined.

Here, we used genome-wide CRISPR screens and cell-free *Xenopus* egg extracts that faithfully model replication-coupled ICL repair to uncover a role for the ATM checkpoint kinase in regulating reversed fork processing during ICL repair. We show that protein phosphatase inhibition disrupts ICL repair by the FA pathway by inducing aberrant

resection of reversed fork intermediates and that this aberrant processing is suppressed by ATM inhibition. We further show that ATM is normally activated upon replication fork reversal and promotes resection of nascent DNA strands, which depends on DNA2, WRN, and EXO1. Our results support a model in which the regressed arm of a reversed fork functions as a one-ended DSB that activates ATM and stimulates DNA2- and EXO1-dependent resection of the regressed arm, thereby enabling fork restoration and completion of ICL repair.

3.3 Results

3.3.1 Chemogenomic screens indicate a role for PP2A during ICL repair by the FA pathway

To identify novel factors that participate in ICL repair by either the FA pathway or alternative repair pathways, we performed genome-wide chemogenomic CRISPR dropout screens with human hTERT-immortalized, p53-deficient RPE-1 cells (Fig. 3.1b and Supplementary Fig. 3.1a,b)[29]. Cells were serially challenged with either UVA (control arm) or psoralen and UVA (PUVA; treatment arm), which together induce the formation of psoralen-ICLs that are known to be repaired by both the FA pathway and an alternative pathway mediated by the NEIL3 glycosylase[7,8,30]. Importantly, we performed screens using both FA pathway-proficient cells (*FANCD2*$^{+/+}$) and isogenic FA pathway-deficient cells (*FANCD2*$^{-/-}$). This approach enabled us to distinguish genes that function epistatically with the FA pathway (only conferring PUVA resistance in the FA pathway-proficient cells) from those that might function in alternative ICL repair pathways (conferring PUVA resistance in both FANCD2-proficient and -deficient cells). Validating our approach, we recovered 16 known *FANC* genes as ICL resistance genes in FANCD2-proficient cells (Fig. 3.1c).

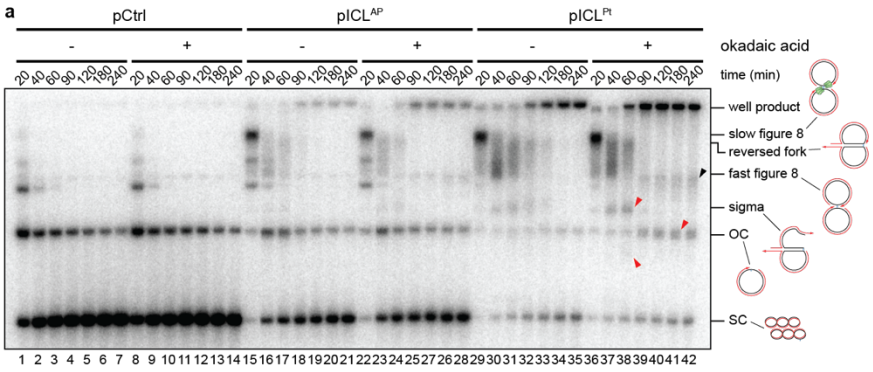
Importantly, except for *FANCJ*, none of these genes were recovered in FANCD2-deficient cells, consistent with roles for these genes in the same ICL repair pathway as *FANCD2*. Conversely, *NEIL3* and *FAN1*, both of which act independently of the FA pathway to promote ICL repair[3,30-32], were recovered as ICL resistance genes in both FANCD2-proficient and -deficient cells.

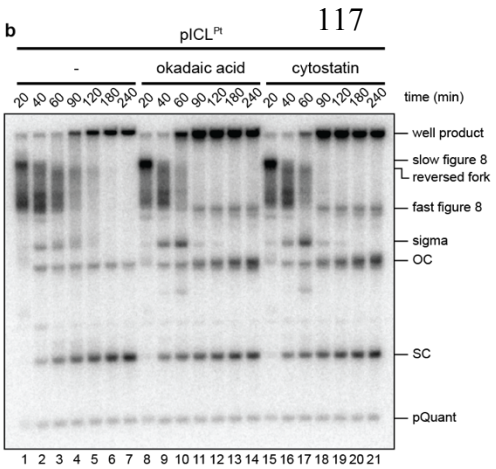
In addition to known *FANC* genes, *PTPA* emerged as an ICL resistance gene specifically in FANCD2-proficient cells. *PTPA* encodes a factor that promotes PP2A phosphatase holoenzyme formation through activation of the PP2A catalytic subunit[33], implying that PP2A promotes ICL repair by the FA pathway. Note that genes encoding PP2A subunits (e.g. PP2A-A and PP2A-C) are essential and would therefore be unlikely to emerge as hits in CRISPR dropout screens[34]. Consistently, another recent set of genome-wide screens also implicated *PTPA* in resistance to cisplatin, which induces ICLs that require the FA pathway for replication-coupled repair[29]. Additionally, PP2A-dependent FANCD2 dephosphorylation has been shown to be critical for FANCD2 ubiquitylation and activation of the FA pathway[35]. We therefore specifically tested whether *PTPA* contributes to ICL resistance. *PTPA* sgRNA knockout cell pools exhibited increased sensitivity to cisplatin (Fig. 3.1d,e), confirming a role for *PTPA* in cellular tolerance to cisplatin-induced DNA damage. Overall, these results suggest that PTPA activates PP2A to promote ICL repair by the FA pathway.

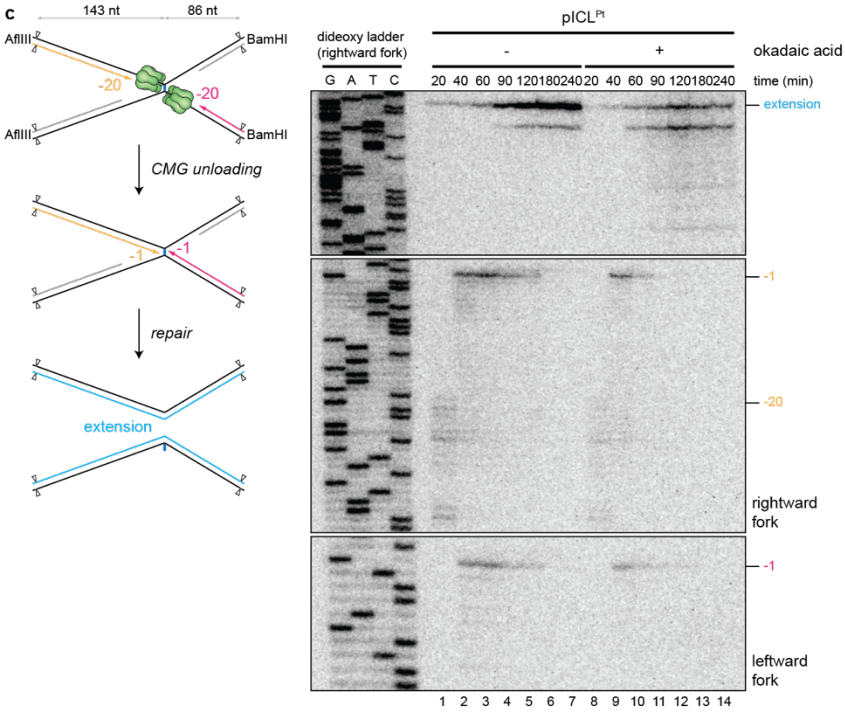
3.3.2 PP2A inhibition destabilizes reversed replication forks

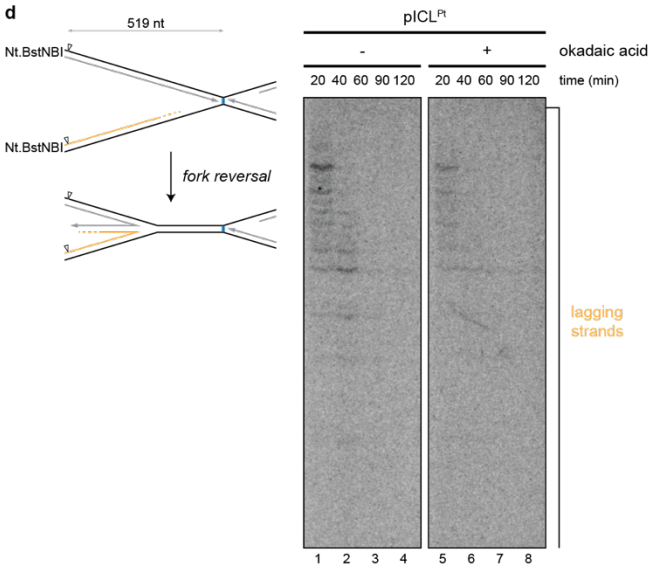
To determine how PP2A phosphatase activity might regulate replication-coupled ICL repair, we replicated undamaged and ICL-containing plasmids in *Xenopus* egg extracts supplemented with the pan-phosphoprotein phosphatase (PPPase) inhibitor okadaic acid.

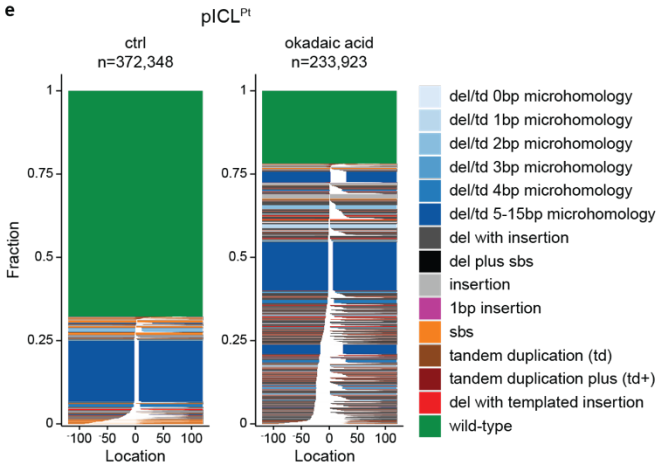
Following replication in the presence of [α-^{32}P]dCTP (which is incorporated into the nascent DNA strands), replication intermediates and products were deproteinized and resolved by native agarose gel electrophoresis (Fig. 3.2a). Since okadaic acid is known to block replication initiation by inhibiting incorporation of CDC45 into pre-replication complexes[36], okadaic acid was added 5 minutes after initiating replication. Addition of okadaic acid did not appreciably affect replication of either an undamaged plasmid (pCtrl) or a plasmid containing an ICL formed by an abasic site (pICL^{AP}), which is predominantly repaired by the NEIL3 pathway in egg extracts. In contrast, okadaic acid had a dramatic impact on replication of a plasmid containing a cisplatin-ICL (pICL^{Pt}) that is repaired by the FA pathway[37]. In the absence of okadaic acid, pICL^{Pt} replication resulted in an initial accumulation of slowly-migrating figure 8 intermediates that correspond to the convergence of replication forks on either side of the ICL. These species were then converted into a topologically distinct, faster-migrating figure 8 intermediate that forms upon unloading of stalled CMG helicases from the ICL. Fast figure 8 intermediates were in turn converted into species with intermediate mobilities that correspond to structures in which one of the replication forks abutting the ICL has undergone extensive reversal with resection of the nascent lagging strand[9]. Subsequent unhooking of the reversed fork intermediates through FANCI-FANCD2-dependent nucleolytic incisions resulted in the formation of sigma products and, eventually, fully-replicated, supercoiled repair products. Strikingly, addition of okadaic acid resulted in a rapid disappearance of reversed fork intermediates and a reduction in the accumulation of fully-replicated and repaired supercoiled products. This was accompanied by a transient but more intense accumulation of sigma products and other faster-migrating species (Fig. 3.2a, red arrowheads)

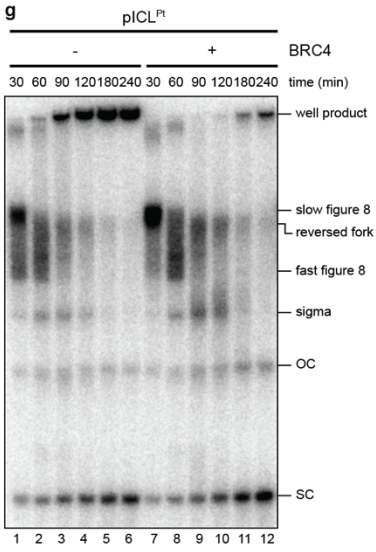

a
pCtrl
pICL^AP
pICL^Pt
okadaic acid
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC


b
pICL^Pt
okadaic acid
cytostatin
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
pQuant


c
143 nt
86 nt
AflIII
BamHI
-20
CMG unloading
-1
repair
extension
dideoxy ladder (rightward fork)
G A T C
pICL^Pt
okadaic acid
time (min)
extension
rightward fork
leftward fork


d
519 nt
Nt.BstNBI
fork reversal
pICL^Pt
okadaic acid
time (min)
lagging strands


e
pICL^Pt
ctrl
n=372,348
okadaic acid
n=233,923
Fraction
Location
del/td 0bp microhomology
del/td 1bp microhomology
del/td 2bp microhomology
del/td 3bp microhomology
del/td 4bp microhomology
del/td 5-15bp microhomology
del with insertion
del plus sbs
insertion
1bp insertion
sbs
tandem duplication (td)
tandem duplication plus (td+)
del with templated insertion
wild-type


g
pICL^Pt
BRC4
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC

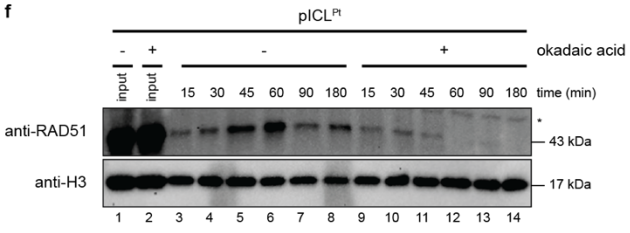

f
pICL^Pt
okadaic acid
input
time (min)
anti-RAD51
anti-H3
43 kDa
17 kDa

Fig. 3.2 | Phosphatase inhibition induces replication fork collapse during ICL repair

a, The indicated plasmids were replicated in egg extract supplemented with [α-^{32}P]dCTP and the PPPase inhibitor okadaic acid, as indicated. Replication intermediates were recovered, deproteinized, separated on a native agarose gel, and visualized by autoradiography. Schematics of replication and repair intermediates that are resolved by native agarose gel electrophoresis are shown. SC, supercoiled; OC, open circular. Red arrowheads indicate fork cleavage products that accumulate during pICL^{Pt} replication in the presence of okadaic acid. Black arrowhead indicates putative fast figure 8 intermediates generated upon reversed fork restoration.

b, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and okadaic acid or cytostatin, as indicated. Replication intermediates were analyzed as in **a**.

c, Left, schematic of nascent leading strands generated during ICL repair. AflIII cuts 143 nt to the left of the cisplatin-ICL. BamHI cuts 86 nt to the right of the cisplatin-ICL. Digestion with AflIII and BamHI generates characteristic -20 to -40 stall, -1 stall, insertion ("0"), and strand extension products. Right, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and okadaic acid, as indicated. Replication intermediates were isolated and digested with AflIII and BamHI, and nascent DNA strands were resolved by denaturing polyacrylamide gel electrophoresis. Top, middle, and bottom panels show sections of the same gels to visualize extension, rightward leading strands, and leftward leading strands, respectively.

d, Left, schematic of nascent lagging strands generated during ICL repair. Nt.BstNBI nicks the rightward nascent lagging strand 519 nt to the left of the cisplatin-ICL. Right, pICL^{Pt} was replicated as in **c**. Replication intermediates were isolated and digested with Nt.BstNBI, and nascent DNA strands were resolved by denaturing polyacrylamide gel electrophoresis.

e, pICL^{Pt} was replicated in egg extract supplemented with okadaic acid, and repair products were PCR amplified and sequenced. Raw paired-end sequencing reads were merged using SIQ, and data were visualized with SIQPlotteR. Color-coded repair outcomes are depicted as tornado plots. White spaces depict deletions. The positions of variations are indicated relative to the cisplatin-ICL site ("0"). n indicates the number of mapped sequencing reads obtained for each condition.

f, pICL^{Pt} was replicated in egg extract supplemented with okadaic acid, as indicated. Recovered chromatin-associated proteins were resolved by SDS-PAGE and blotted for RAD51 and H3 (loading control).

g, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and BRC4 peptide, as indicated. Replication intermediates were analyzed as in **a**.

corresponding to replication forks that have undergone cleavage of one or both parental DNA strands (Supplementary Fig. 3.2a,b). Okadaic acid also induced an earlier and more pronounced accumulation of well products, which may indicate the formation of end-joining intermediates and products following okadaic acid-induced fork cleavage[38]. Additionally, we observed the reappearance and persistence of species that co-migrate with fast figure 8 intermediates (Fig. 3.2a, black arrowhead), suggestive of premature reversed fork restoration prior to the occurrence of unhooking incisions. These observations suggest that PPPase activity promotes ICL repair by the FA pathway by ensuring proper processing of reversed fork intermediates. Treatment with the PP2A-selective inhibitor cytostatin[39] similarly resulted in destabilization of pICL^{Pt} reversed fork intermediates and formation of fork cleavage products, well products, and putative cross-linked fork restoration products (Fig. 3.2b), indicating that PP2A, specifically, is required for efficient ICL repair by the FA pathway. Further, addition of excess recombinant human PP2A-AC subcomplex largely suppressed the effects of okadaic acid on pICL^{Pt} replication (Supplementary Fig. 3.2c-e).

To better understand how okadaic acid disrupts ICL repair by the FA pathway, we monitored the stability and maturation of the nascent DNA strands during ICL repair. pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and okadaic acid, as indicated, and recovered replication intermediates were digested with restriction enzymes that allow the nascent DNA leading strands to be resolved on a denaturing sequencing gel (Fig. 3.2c). In the absence of okadaic acid, nascent leading strands initially stalled 20 to 40 nucleotides (nt) upstream of the ICL (-20 to -40 stall) due to the footprint of CMG. Upon

CMG unloading, the nascent strands were extended up to 1 nt upstream of the ICL (-1 stall). ICL unhooking and TLS past the resulting ICL remnant converted the nascent leading strands into full-length extension products. In the presence of okadaic acid, nascent leading strands first stalled at -20 to -40 before being extended to the -1 position. However, okadaic acid treatment induced a rapid disappearance of nascent leading strand -1 stall products. This loss of leading strand stall products was accompanied by both a decrease in formation of expected full-length nascent strand extension products and an increase in nascent strand extension products that were predominantly shorter than the expected product, suggesting that excessive okadaic acid-induced processing of replication intermediates leads to fork collapse and formation of deletion-prone end-joining products. We also specifically examined the fate of nascent lagging strands using a nicking enzyme that liberates the lagging strand from one fork. Okadaic acid treatment accelerated the disappearance of nascent lagging strands (Fig. 3.2d; compare lanes 1 and 2 with lanes 5 and 6), consistent with accelerated 5' to 3' resection of reversed fork intermediates. Sequencing of $pICL^{Pt}$ replication products recovered from untreated or okadaic acid-treated extracts confirmed that okadaic acid treatment caused a reduction in the fraction of sequencing reads corresponding to error-free repair products and a substantial increase in the fraction of reads with deletions spanning the ICL (Fig. 3.2e). Taken together, our results support a model in which PP2A limits resection of repair intermediates, which can lead to fork collapse and error-prone end-joining.

PP2A has been reported to physically interact with BRCA2 and promote RAD51 loading at DSBs[40]. Since RAD51 filaments antagonize reversed fork degradation by DNA2 and MRN, we considered the possibility that excessive reversed fork intermediate resection in

okadaic acid-treated extracts results from inefficient RAD51 loading and a consequent defect in fork protection. Consistent with this model, okadaic acid reduced the amount of RAD51 associated with chromatin recovered from pICL^{Pt} replication reactions (Fig. 3.2f). We next tested whether disruption of RAD51 loading is sufficient to destabilize reversed fork intermediates formed during replication-coupled ICL repair. As previously observed[13], addition of BRC4 peptide that blocks the interaction between RAD51 and BRCA2 prevented RAD51 loading during replication of pICL^{Pt} (Supplementary Fig. 3.2f). However, addition of BRC4 peptide did not induce excessive resection of reversed fork intermediates (Fig. 3.2g), indicating that loss of RAD51-dependent fork protection is not sufficient to induce reversed fork degradation during ICL repair. We infer that, although okadaic acid may compromise RAD51-dependent fork protection, increased resection of reversed fork intermediates in okadaic-acid treated extracts occurs by another mechanism, such as hyperactivation of nucleases that mediate resection.

3.3.3 ATM promotes reversed fork processing during ICL repair

PP2A targets thousands of phosphosites across the human proteome[41], making the identification of specific dephosphorylation events that regulate ICL repair a formidable challenge. However, we reasoned that PP2A activity would be likely to counteract one or more protein kinases that are activated during ICL repair. Accordingly, inhibition of relevant kinase activities should prevent excessive resection of reversed fork intermediates during ICL repair in phosphatase-inhibited egg extract. We therefore tested whether chemical inhibition of CDKs or the apical DNA damage response kinases ATR and ATM suppresses the effects of okadaic acid on repair intermediate processing during pICL^{Pt} replication in egg extract. Pan-CDK inhibition with flavopiridol did not noticeably affect

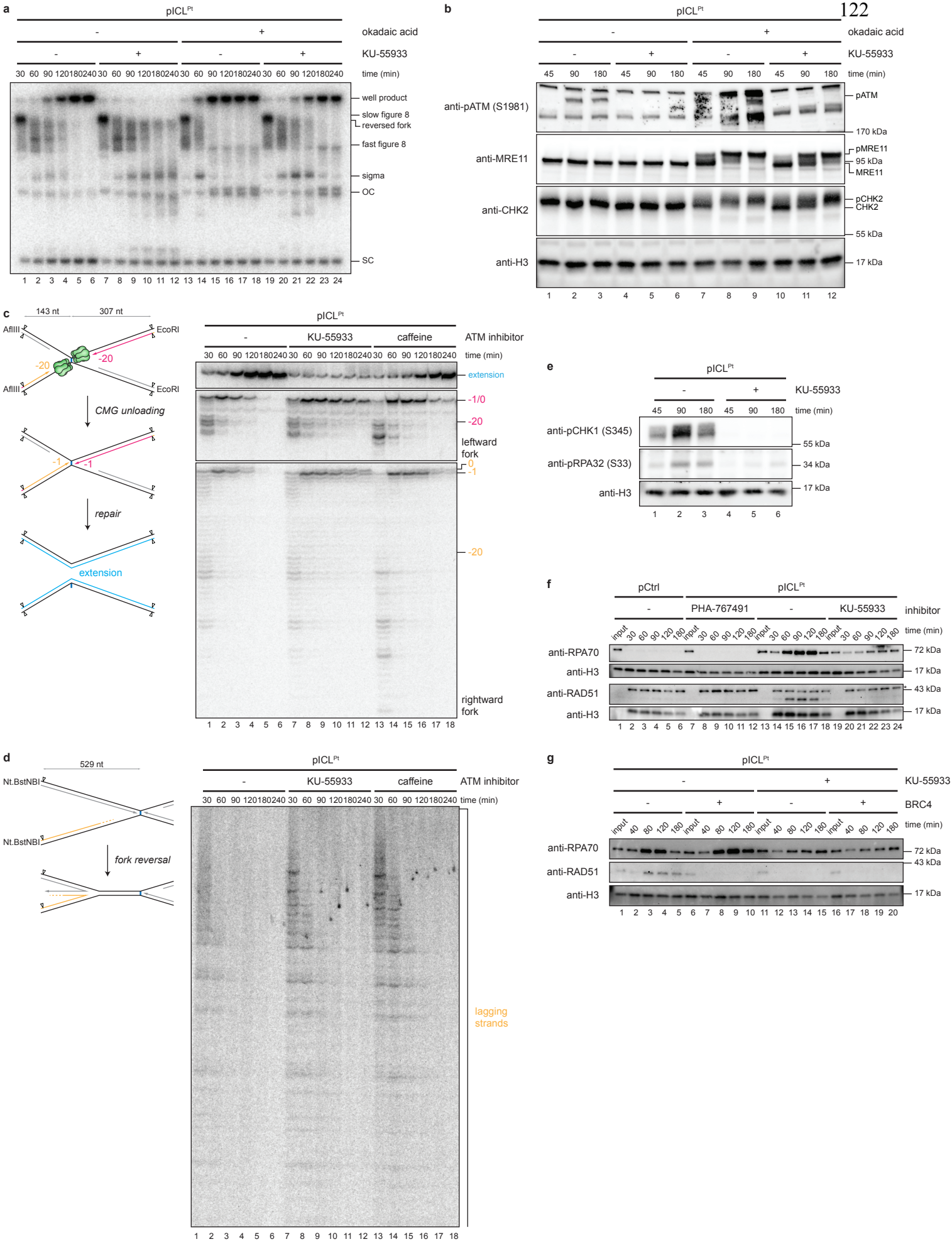

a
pICL^Pt
okadaic acid
KU-55933
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
b
pICL^Pt
okadaic acid
KU-55933
time (min)
anti-pATM (S1981)
pATM
170 kDa
anti-MRE11
pMRE11
95 kDa
MRE11
anti-CHK2
pCHK2
CHK2
55 kDa
anti-H3
17 kDa
c
143 nt
307 nt
AflIII
EcoRI
CMG unloading
repair
extension
pICL^Pt
ATM inhibitor
KU-55933
caffeine
time (min)
extension
-1/0
-20
leftward fork
rightward fork
d
529 nt
Nt.BstNBI
fork reversal
pICL^Pt
ATM inhibitor
KU-55933
caffeine
time (min)
lagging strands
e
pICL^Pt
KU-55933
time (min)
anti-pCHK1 (S345)
anti-pRPA32 (S33)
anti-H3
55 kDa
34 kDa
17 kDa
f
pCtrl
pICL^Pt
PHA-767491
KU-55933
inhibitor
time (min)
anti-RPA70
anti-H3
anti-RAD51
72 kDa
43 kDa
17 kDa
g
pICL^Pt
KU-55933
BRC4
time (min)
anti-RPA70
anti-RAD51
anti-H3
72 kDa
43 kDa
17 kDa

Fig. 3.3 | ATM promotes reversed fork processing during ICL repair

a, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and okadaic acid and the ATM inhibitor KU-55933, as indicated. Replication intermediates were analyzed as in Fig. 3.2a.

b, pICL^{Pt} was replicated in egg extract supplemented with okadaic acid and KU-55933, as indicated. Replication reactions were separated by SDS-PAGE and blotted for phospho-ATM (S1981), MRE11, CHK2, and H3 (loading control).

c, Nascent DNA leading strands from pICL^{Pt} replication reactions supplemented with [α-^{32}P]dCTP and KU-55933 or caffeine were digested with AflIII and EcoRI and analyzed as in Fig. 3.2c. AflIII cuts 143 nt to the left of the cisplatin-ICL. EcoRI cuts 307 nt to the right of the cisplatin-ICL.

d, Nascent DNA lagging strands from pICL^{Pt} replication reactions supplemented with [α-^{32}P]dCTP and KU-55933 or caffeine were digested with Nt.BstNBI and analyzed as in Fig. 3.2d.

e, pICL^{Pt} was replicated in egg extract supplemented with KU-55933, as indicated. Replication reactions were resolved and blotted for phospho-CHK1 (S345), phospho-RPA32 (S33) and H3 (loading control) as in **b**.

f, pICL^{Pt} was replicated in egg extract supplemented with PHA-767491 (CDC7 inhibitor) and KU-55933, as indicated. Recovered chromatin-associated proteins were blotted for RPA70, RAD51, and H3 (loading control) as in Fig. 3.2f. Asterisk indicates non-specific band.

g, pICL^{Pt} was replicated in egg extract supplemented with BRC4 peptide and KU-55933, as indicated. Recovered chromatin-associated proteins were blotted for RPA70, RAD51, and H3 (loading control) as in Fig. 3.2f.

the formation and resolution of pICL^{Pt} repair intermediates and did not suppress okadaic acid-induced destabilization of reversed fork intermediates (Supplementary Fig. 3.3a,b). On its own, addition of the ATR inhibitor AZD6738 caused a subtle accumulation of reversed fork intermediates (Supplementary Fig. 3.3c), consistent with a role for ATR in stimulating both unhooking by the FA pathway and DSB end resection[1,42]. ATR inhibition also partially stabilized reversed fork intermediates and partially suppressed the accumulation of putative fork cleavage products in okadaic acid-treated extract

(Supplementary Fig. 3.3c), indicating that ATR may stimulate resection of reversed forks during ICL repair.

We then examine the effects of inhibiting ATM, which is generally associated with processing of DSB ends but not with processing of replication intermediates. Surprisingly, addition of the ATM inhibitor KU-55933 during pICL^{Pt} replication resulted in a strong accumulation of reversed fork intermediates (Fig. 3.3a). KU-55933 addition also increased the stability of reversed fork intermediates and delayed the formation of fork cleavage products during pICL^{Pt} replication in okadaic acid-treated extract, suggesting that PP2A counteracts ATM to regulate processing of reversed fork intermediates during ICL repair. We observed similar stabilization of reversed forks – both in the presence and absence of okadaic acid – in extracts supplemented with another ATM inhibitor (AZD0156) as well as with caffeine, an inhibitor of both ATM and ATR (Supplementary Fig. 3.3d,e). Consistent with the idea that PP2A counteracts ATM activation, okadaic acid treatment during pICL^{Pt} replication induced increased phosphorylation of ATM at S1981, a canonical marker of ATM activation[43,44], as well as phosphorylation of the ATM target proteins MRE11 and CHK2 (Fig. 3.3b). This phosphorylation of ATM, MRE11, and CHK2 was attenuated by addition of ATM inhibitor. Combined, our results indicate that PP2A prevents hyperactivation of ATM during ICL repair. This conclusion aligns with previous work demonstrating that PP2A interacts with ATM and suppresses ATM activation in the absence of DNA damage[45].

We then determined how ATM inhibition influences the processing and maturation of nascent DNA strands. ATM inhibition with either KU-55933 or caffeine caused a strong accumulation of nascent leading strand -1 and 0 stall products and a corresponding

reduction in full length extension products (Fig. 3.3c). ATM inhibition also resulted in stabilization of nascent lagging strands (Fig. 3.3d). Combined, these results suggest that ATM stimulates resection of reversed forks. To further test whether ATM is required for resection of reversed fork intermediates, we examined the extent of CHK1 and RPA phosphorylation during pICL^{Pt} replication. CHK1 and RPA are phosphorylated by ATR upon accumulation of single stranded (ss)DNA and serve as markers of DNA resection[46]. ATM inhibition during ICL repair suppressed the accumulation of phosphorylated CHK1 and RPA in extract (Fig. 3.3e), providing further evidence that ATM stimulates resection. Furthermore, ATM inhibition abolished replication- and damage-dependent loading of RAD51 onto chromatin and attenuated the loading of RPA, consistent with defective resection (Fig. 3.3f and Supplementary Fig. 3.3f). Since ATM can also promote RAD51 loading by phosphorylating factors that mediate RAD51 loading, including BRCA2 and PALB2[24,40], we wondered whether the failure of RAD51 to bind chromatin in ATM-inhibited extract reflects a deficit in RAD51 mediator function at sites of ssDNA rather than a failure to perform resection. We therefore compared the association of RPA and RAD51 with chromatin during ICL repair in extract that was treated with either BRC4 peptide or ATM inhibitor. BRC4 peptide blocked RAD51 association with chromatin and caused an accumulation of RPA on chromatin (Fig. 3.3g and Supplementary Fig. 3.3g), indicative of impaired exchange of RPA for RAD51 on ssDNA. By contrast, ATM inhibition impaired the association of both RAD51 and RPA with chromatin (Fig. 3.3g), consistent with a deficit in resection and a consequent reduction in ssDNA. In aggregate, our data indicate that ATM promotes resection of reversed fork intermediates during ICL repair.

3.3.4 ATM is activated concomitantly with fork reversal

We next tested whether ATM becomes activated by a specific DNA structure during ICL repair. pCtrl and pICL^{Pt} were replicated in egg extracts, and ATM activation was monitored by immunoblotting for ATM phosphorylation at S1981. Whereas ATM phosphorylation was undetectable during mock replication in the absence of plasmid as well as during replication of undamaged pCtrl, phosphorylated ATM accumulated in a time-dependent manner during pICL^{Pt} replication (Fig. 3.4a and Supplementary Fig. 3.4a), indicating that ATM activation is damage-dependent. During ICL repair by the FA pathway, fork reversal produces a regressed DNA arm that is equivalent to a one-ended DSB, which we suspected might trigger ATM activation. Consistent with this hypothesis, ATM phosphorylation was not activated during pICL^{Pt} replication in extract supplemented with p97 inhibitor (Fig. 3.4a and Supplementary Fig. 3.4a), which blocks CMG unloading and subsequent fork reversal. A DSB is also generated through FANCI- and FANCD2-dependent incisions that unhook the ICL[12]. We therefore examined the extent of ATM activation during pICL^{Pt} replication in FANCD2-depleted extract. FANCD2 depletion led to a pronounced reduction in the amount of sigma intermediate incision products (Supplementary Fig. 3.4b,c) but did not diminish the extent of ATM phosphorylation (Fig. 3.4b), indicating that the reversed fork structure itself can activate ATM. We therefore conclude that the reversed fork structure generated during ICL repair is sufficient to activate ATM and processing of the regressed arm.

We also determined whether ATM hyperactivation can induce processing of replication intermediates other than reversed fork structures. pICL^{Pt} was replicated in egg extract that was supplemented with okadaic acid and p97 inhibitor, as indicated, and replication

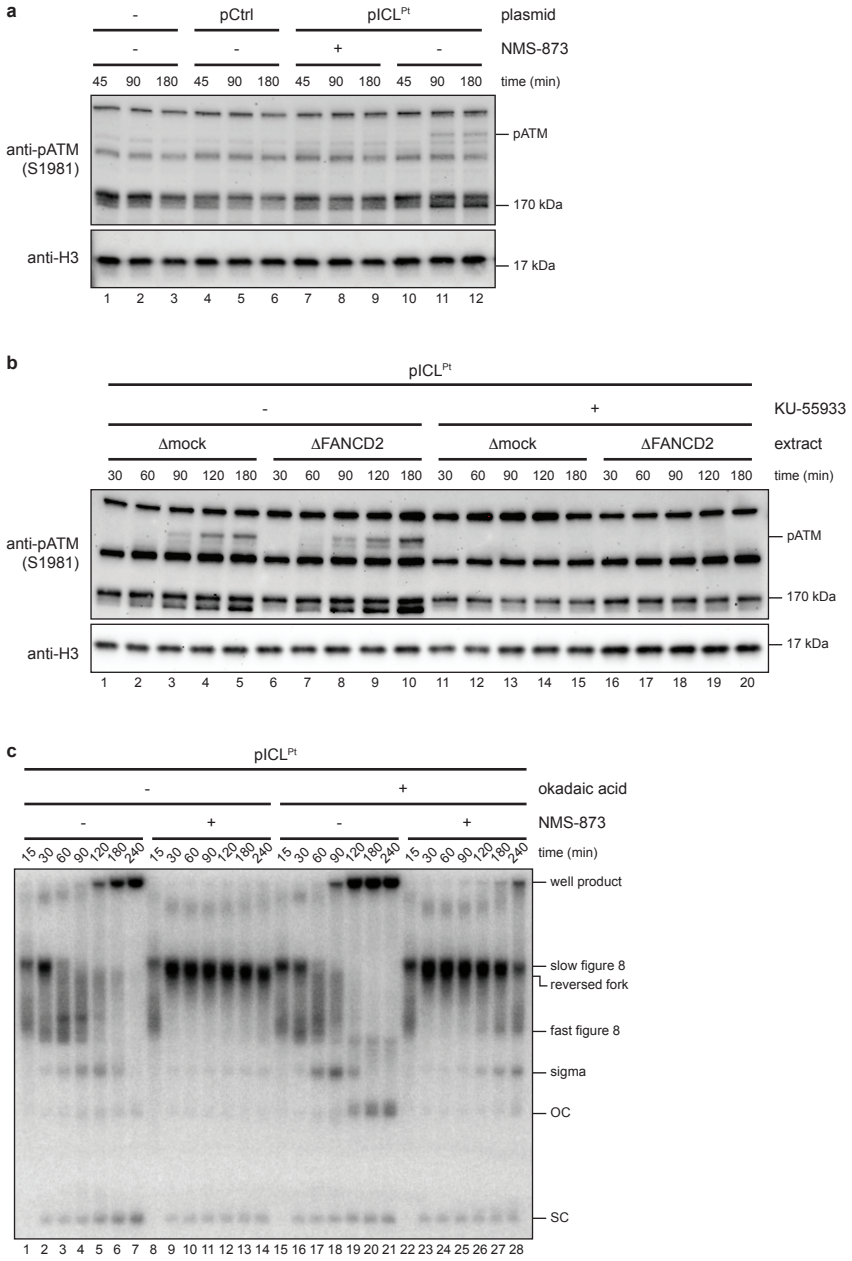


Fig. 3.4 | ATM activation during ICL repair requires CMG unloading

a, Plasmids were replicated in egg extract supplemented with NMS-873 (p97 inhibitor), as indicated. Replication reactions were blotted for phospho-ATM (S1981) and H3 (loading control) as in Fig. 3.3b.

b, pICL^{Pt} was replicated in mock- or FANCD2-depleted egg extract supplemented with KU-55933, as indicated. Replication reactions were blotted for phospho-ATM (S1981) and H3 (loading control) as in Fig. 3.3b.

c, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and okadaic acid and NMS-873, as indicated. Replication intermediates were analyzed as in Fig. 3.2a.

intermediates were resolved on a native agarose gel (Fig. 3.4c). As expected, in the absence of okadaic acid, p97 inhibitor blocked the formation of reversed fork intermediates and caused a persistence of slow figure 8 intermediates that form when forks converge at the ICL. Importantly, these slow figure 8 intermediates also persisted in extract supplemented with both okadaic acid and p97 inhibitor, indicating that these structures are resistant to ATM-induced processing (Fig. 3.4c). The slow figure 8 intermediates remained stable in okadaic acid-treated extracts in which ATM activation was further stimulated by addition of linear plasmid DNA (Supplementary Fig. 3.4d,e). These data suggest that ATM-dependent nucleolytic processing is specifically targeted to reversed fork intermediates.

3.3.5 DNA2, WRN, and EXO1 mediate reversed fork resection during ICL repair

To identify potential effectors of ATM-dependent reversed fork resection, we profiled the phosphoproteome during $pICL^{Pt}$ replication in untreated and okadaic acid-treated extracts. Okadaic acid treatment resulted in a significant, at least 2-fold increase in abundance for ~8.7% of phosphopeptides detected (Fig. 3.5a and Supplementary Fig. 3.5a,b). Several phosphopeptides that were significantly enriched in okadaic acid-treated extract corresponded to the nucleases DNA2 and EXO1 as well as to the DNA2-associated helicases BLM and WRN. Since DNA2, WRN, and EXO1 have been previously implicated in reversed fork processing[18,47], we tested whether they mediate resection during ICL repair. Addition of the DNA2 inhibitor C5[48], the EXO1 inhibitor C200[49], or the WRN inhibitor HRO761[50] to extracts resulted in a strong accumulation of reversed fork intermediates and a decrease in the amount of fully-replicated supercoiled products (Fig. 3.5b), mirroring the effects of ATM inhibition and suggesting that DNA2, WRN, and EXO1 process the reversed fork intermediates and enable completion of ICL repair. Similar

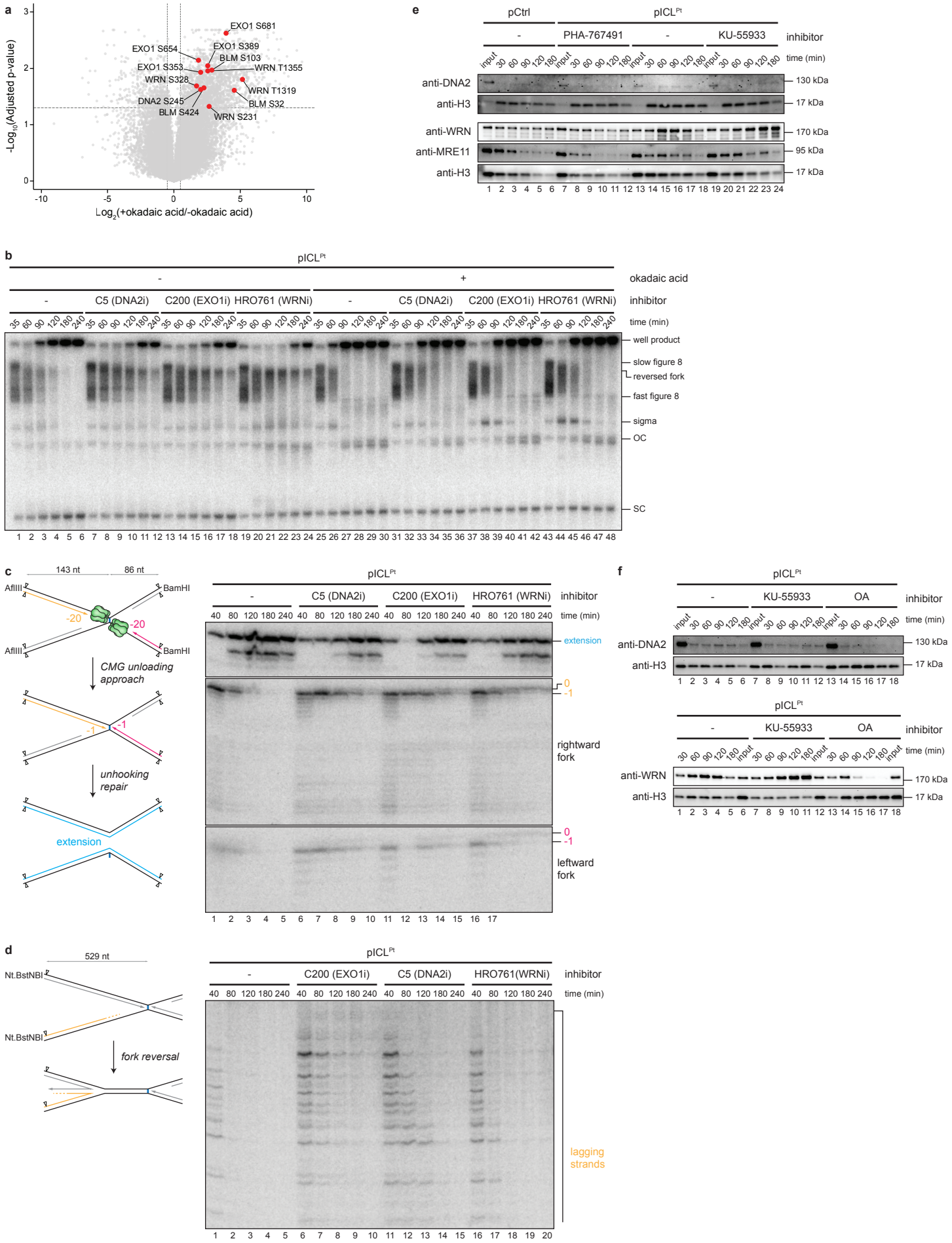

a
EXO1 S681
EXO1 S654
EXO1 S389
BLM S103
EXO1 S353
WRN T1355
WRN S328
WRN T1319
DNA2 S245
BLM S32
BLM S424
WRN S231
-Log10(Adjusted p-value)
Log2(+okadaic acid/-okadaic acid)
e
pCtrl
pICLPt
PHA-767491
KU-55933
inhibitor
time (min)
anti-DNA2
anti-H3
anti-WRN
anti-MRE11
130 kDa
17 kDa
170 kDa
95 kDa
b
okadaic acid
C5 (DNA2i)
C200 (EXO1i)
HRO761 (WRNi)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
c
143 nt
86 nt
AflIII
BamHI
-20
CMG unloading approach
-1
unhooking repair
extension
C5 (DNA2i)
C200 (EXO1i)
HRO761 (WRNi)
rightward fork
leftward fork
d
529 nt
Nt.BstNBI
fork reversal
C200 (EXO1i)
C5 (DNA2i)
HRO761(WRNi)
lagging strands
f
KU-55933
OA
anti-DNA2
anti-H3
anti-WRN

Fig. 3.5 | EXO1 and DNA2-WRN mediate processing of reversed fork intermediates during ICL repair

a, Volcano plots comparing the abundance of phosphopeptides 90 minutes after initiating pICL^{Pt} replication in untreated and okadaic acid-supplemented egg extract. The estimated fold change (FC) in normalized phosphopeptide intensity between the samples is plotted against the false discovery rate- (FDR-)adjusted p-value. Labeled proteins of interest meet significance criteria ($|Log_2FC| > 0.5$ and $p < 0.05$).

b, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and okadaic acid and the DNA2 inhibitor C5, the EXO1 inhibitor C200, or the WRN inhibitor HRO761, as indicated. Replication intermediates were analyzed as in Fig. 3.2a.

c, Nascent DNA leading strands from pICL^{Pt} replication reactions supplemented with [α-^{32}P]dCTP and C5, C200, or HRO761 were digested with AflIII and BamHI and analyzed as in Fig. 3.2c.

d, Nascent DNA lagging strands from pICL^{Pt} replication reactions supplemented with [α-^{32}P]dCTP and C200, C5, or HRO761 were digested with Nt.BstNBI and analyzed as in Fig. 3.2d.

e, pCtrl and pICL^{Pt} were replicated in egg extract supplemented with PHA-767491 or KU-55933. Recovered chromatin-associated proteins were blotted for DNA2, WRN, MRE11, and H3 (loading control) as in Fig. 3.2f.

f, pICL^{Pt} was replicated in egg extract supplemented with KU-55933 or okadaic acid. Recovered chromatin-associated proteins were blotted for DNA2, WRN, and H3 (loading control) as in Fig. 3.2f.

stabilization of reversed fork intermediates generated during ICL repair was also observed with the EXO1 inhibitor F684[49] (Supplementary Fig. 3.5c). Inhibition of DNA2, WRN, or EXO1 also stabilized reversed fork intermediates and delayed the accumulation of putative fork collapse products in okadaic acid-treated extract (Fig. 3.5b), implying that these nucleases and helicases drive reversed fork over-resection in the absence of PP2A activity.

We next examined the effects of DNA2, WRN, and EXO1 inhibition on nascent strand maturation. Like ATM inhibition, inhibition of DNA2, WRN, or EXO1 caused a strong persistence of -1 and 0 leading strand stall products and a corresponding reduction in the amount of full-length extension products (Fig. 3.5c), suggesting that DNA2, WRN, and

EXO1 might be required for reversed fork restoration and subsequent bypass of the unhooked ICL remnant. Inhibition of DNA2, WRN, or EXO1 also caused a strong persistence of nascent lagging strands, indicating that these nucleases promote 5' to 3' resection of reversed forks (Fig. 3.5d). Interestingly, EXO1 inhibition resulted in a more pronounced stabilization of the longest lagging strands with 5' ends nearest to the end of the regressed arm of a reversed fork structure. By contrast, inhibition of DNA2 or WRN resulted in more pronounced stabilization of shorter lagging strands corresponding to recessed 5' ends on the regressed arm of a reversed fork. These data suggest that EXO1 may initiate 5' to 3' resection at a reversed fork, after which DNA2 acts with WRN to further resect 5' recessed ends.

To determine how ATM influences resection factor dynamics during ICL repair, we replicated $pICL^{Pt}$ in egg extract and examined the association of DNA2, WRN, and the MRN nuclease subunit MRE11 with recovered chromatin. DNA2 and WRN exhibited replication- and damage-dependent recruitment to chromatin (Fig. 3.5e and Supplementary Fig. 3.5d). ATM inhibition caused a delay in the association of DNA2 with chromatin, implying that ATM promotes DNA2 recruitment during ICL repair (Fig. 3.5e). The effect of ATM inhibition on DNA2 recruitment was most evident 60-120 minutes after initiating replication, when reversed fork intermediates first accumulate (Supplementary Fig. 3.5d). ATM inhibition also delayed the accumulation of WRN on chromatin (Fig. 3.5e), suggesting that WRN recruitment depends on ATM activity. Notably, phosphatase inhibition with okadaic acid led to a rapid departure of DNA2 and WRN from chromatin (Fig. 3.5f), consistent with accelerated resection of reversed fork intermediates in phosphatase-inhibited extract (Supplementary Fig. 3.5e). Like DNA2 and WRN, MRE11

accumulated on chromatin during ICL repair (Fig. 3.5e). However, our data do not distinguish whether ATM enhances MRN recruitment to chromatin or whether MRN nuclease activity contributes to resection of reversed fork intermediates during ICL repair. Importantly, accumulation of MRE11 on chromatin during ICL repair suggests that MRN may mediate recruitment and activation of ATM at the regressed arm of the reversed fork, paralleling MRN-dependent activation of ATM at DSBs[22]. Taken together, our data indicate that fork reversal promotes activation of ATM, which in turn stimulates resection, at least in part, by facilitating nuclease recruitment.

3.4 Discussion

Although current data indicate that replication fork reversal and restoration are integral steps in ICL repair by the FA pathway, little is known about which factors mediate these fork dynamics or how they are regulated. Here, we used chemogenomic screens and *Xenopus* egg extracts to uncover a phosphoregulatory circuit involving ATM that contributes to reversed fork processing during ICL repair. A CRISPR dropout screen implicated the PP2A activator PTPA in ICL repair by the FA pathway. We then determined that inhibition of protein phosphatase activity results in hyperactivation of ATM and is associated with destabilization of reversed fork intermediates, fork collapse, and formation of deletion-prone end-joining products. Our data indicate that ATM is activated concurrently with fork reversal and that this activation is required for 5’ to 3’ resection of nascent lagging strands (Fig. 3.6a). These observations indicate that ATM directs resection of reversed forks in a manner that is counteracted by phosphatase activity. We further implicate the long-range resection enzymes DNA2-WRN and EXO1 in mediating this resection, implying that these factors are directly or indirectly activated by ATM at a

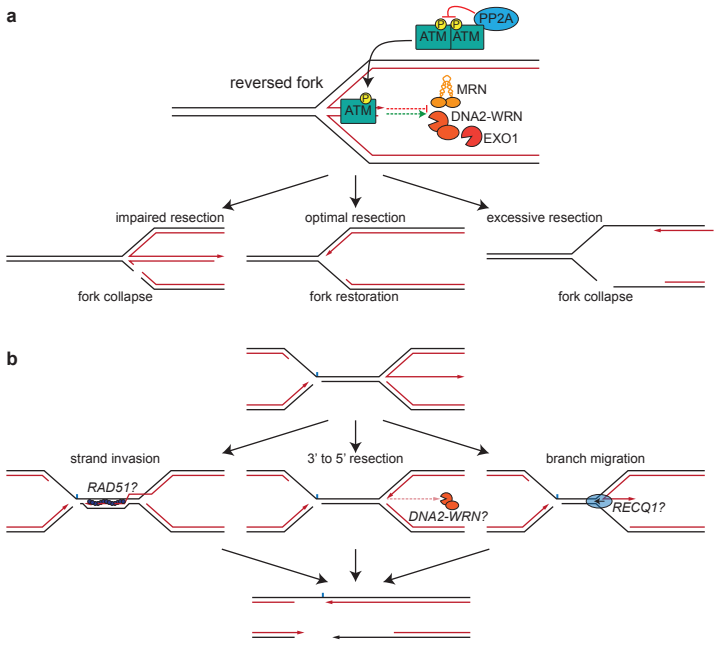


Fig. 3.6 | Model for regulation of reversed fork processing by ATM

See main text for details.

a, Fork reversal generates a regressed arm that activates ATM. ATM activation directly or indirectly stimulates nascent strand resection by EXO1 and DNA2-WRN while preventing excessive resection by MRN. PP2A suppresses ATM activation to constrain resection.

b, Following fork reversal and ICL unhooking by the FA pathway, fork restoration enables completion of replication and repair. Potential mechanisms by which resection could promote fork restoration are shown.

reversed fork. Thus, our results indicate that the reversed fork formed during ICL repair is recognized as a DSB and initiates a checkpoint signaling cascade. Given that replication fork reversal is a frequent response to diverse replication stresses[51], ATM may be a key regulator of the replication stress response.

Undoubtedly, ATM has pleiotropic effects on ICL repair and processing of reversed forks. At DSB ends, ATM is known to phosphorylate EXO1 and to physically interact with DNA2[52,53]. ATM also phosphorylates the WRN and BLM helicases, both of which stimulate DNA2-dependent resection and have been implicated in the processing of reversed forks[18,54-59]. Our data therefore indicate that ATM may directly activate DNA2,

WRN, and/or EXO1 to stimulate resection at reversed forks, either by stimulating the activity of these enzymes or by promoting their recruitment to reversed forks. ATM may also contribute to reversed fork resection by more indirect mechanisms, such as by enabling displacement of the anti-resection factor 53BP1 from chromatin[60,61]. Conversely, ATM has also been reported to stabilize reversed forks. MRN phosphorylation by ATM facilitates MRN release from chromatin and prevents the degradation of reversed forks[25]. Combined with this previous report, our results indicate that ATM regulates multiple stages of resection to tune reversed fork processing, promoting sufficient resection for replication restart while avoiding excessive resection and consequent fork collapse (Fig. 3.6a).

PP2A also appears to regulate multiple facets of ICL repair. PP2A enables efficient ubiquitylation and chromatin recruitment of the FANCI-FANCD2 complex at ICLs[35]. PP2A also interacts with BRCA2 to promote RAD51 filament formation during homologous recombination[40], suggesting that it may contribute to fork protection during ICL repair or facilitate repair of the DSB intermediate generated by FANCI-FANCD2-dependent ICL unhooking. Our data now reveal a distinct function for PP2A that contributes to efficient ICL repair. By counteracting ATM activation, PP2A prevents excessive resection of reversed fork intermediates. Notably, while reversed forks underwent rapid degradation upon PP2A inhibition, reversed forks were not destabilized by inactivation of either FANCD2 or RAD51 (Fig. 3.2b,g and Supplementary Fig. 3.4c), both of which contribute to fork protection during ICL repair[62,63]. Thus, in egg extracts, PP2A may preserve fork integrity primarily by suppressing ATM-stimulated resection rather than by enhancing fork protection pathways. These findings further indicate that multiple phosphatases cooperate to limit resection at reversed forks. The RIF1-PP1

complex, for example, dephosphorylates WRN and suppresses DNA2-WRN-dependent fork degradation[64,65]. We propose that PP2A-mediated ATM dephosphorylation and RIF1-PP1-mediated dephosphorylation of resection enzymes may act in concert to restrain pathological fork resection during ICL repair, possibly alongside additional phosphatase activities.

A key outstanding question concerns how ATM-stimulated reversed fork resection might promote fork restoration and the completion of ICL repair. We can envision multiple possibilities. 5' to 3' resection of the nascent lagging strand could allow RAD51 to load onto the resulting 3' overhang and mediate invasion of the parental DNA duplex, generating a joint molecule structure with the nascent leading strand poised for extension past the unhooked ICL remnant (Fig. 3.6b, left). While this idea may be consistent with available data, evidence for the formation of strand invasion intermediates predicted by this model is currently lacking[9]. 5' to 3' resection could also act in concert with a 3' to 5' exonuclease or 3' flap endonuclease that degrades the 3' overhang, eventually regenerating a replication fork with a free 3' end that can be extended through the parental DNA duplex and past the ICL remnant (Fig. 3.6b, middle). Interestingly, our data indicate that nascent leading strand -1 stall products are highly unstable in extract treated with phosphatase inhibitor, though it is unclear whether these nascent leading strands are processed though end-joining or nucleolytic degradation. Nonetheless, it is worth noting that DNA2, which our data implicate in reversed fork processing during ICL repair, has robust 3' to 5' nuclease activity that can, in some contexts, mediate nascent leading strand degradation at a reversed fork[66,67]. Resection of the regressed arm might also cooperate with branch migration to facilitate fork restoration by a DNA translocase, such as a member of the RECQ family

(Fig. 3.6b, right). If targeted to both strands of the regressed arm, resection would reduce the distance of branch migration required before a replication fork is restored. It is also possible that by converting the regressed arm into a ssDNA 3' flap, resection imparts a polarity at the reversed fork four-way junction that biases branch migration in favor of fork restoration rather than further reversal. ATM-dependent reversed fork processing could also indirectly favor fork restoration by antagonizing toxic non-homologous end-joining of reversed fork regressed arms[68,69]. Alternatively, ATM-stimulated end resection could have a neutral or even deleterious effect on fork restoration. In this case, resection might reflect specious ATM activation at a reversed fork that is normally suppressed by phosphatases.

Our results suggest a mechanistic basis for understanding previously reported genetic interactions between DNA2 and the FA pathway. DNA2 depletion induces hypersensitivity to cisplatin, implying that DNA2-mediated resection is required during ICL repair[70]. Counterintuitively, DNA2 depletion suppresses ICL sensitivity in FANCD2-deficient cells, indicating that DNA2-dependent resection of ICL repair intermediates can corrupt ICL repair when the FANCD2 function is compromised[71]. We propose that DNA2 deficiency can suppress defects in the FA pathway by extending the window during which unhooking can occur before nonproductive fork restoration or excessive fork resection renders the ICL refractory to repair. This model suggests that manipulating the balance between fork reversal and fork restoration could have therapeutic benefit in FA patients. Low doses of ATM inhibitors could stabilize reversed fork intermediates to facilitate ICL unhooking by a compromised FA pathway, though FA-deficient cells are profoundly sensitive to ATM inhibition due to an inability to activate homologous recombination and suppress toxic non-

homologous end-joining[72]. However, it may be possible to dampen reversed fork resection and rescue ICL repair using small molecule activators of PP2A (SMAPs)[73].

3.5 Methods

3.5.1 Oligonucleotides

Oligonucleotides used for preparation of cross-linked inserts and sequencing ladders were ordered from Integrated DNA Technologies, and their sequences can be found in Supplementary Table 3.1.

3.5.2 Cell lines

RPE1-hTERT Cas9 *TP53*$^{-/-}$ were a gift from the Durocher lab and generated as described[74]. RPE1-hTERT Cas9 *TP53*$^{-/-}$ *FANCD2*$^{-/-}$ were generated by nucleofecting RPE1-hTERT Cas9 *TP53*$^{-/-}$ cells with a plasmid containing the sgRNA sequence targeting FANCD2 (5'-ATTCCCAGCACGCTGATGTG-3) using a custom nucleofection buffer (25 mM Na_2HPO_4 [pH 7.75], 2.5 mM KCl, 11 mM $MgCl_2$) and Lonza nucleofection cuvettes in an Amaxa 4D nucleofector using program CN114. 24 hours after transfection, cells were selected with gentamicin (Thermo Fisher Scientific #15-710-064). After an additional 72 hours, single cells were sorted into 96-well plates on a BD FACSARIA II Cell sorter instrument and grown until colonies formed. *FANCD2*$^{-/-}$ clones were selected based on successful gene editing determined by PCR amplification and TIDE analysis[75], and the knockout phenotype was confirmed via immunoblotting for FANCD2. RPE1-hTERT-derived cell lines were grown in Dulbecco's Modified Eagle medium (DMEM; Gibco #11965126) supplemented with 10% heat inactivated fetal bovine serum (Gibco #10082147) and 2 μg/mL blasticidin (Invivogen #ant-bl-05). HEK Lenti-X 293T (Takara

#632180) cells were grown in DMEM (Gibco #10569010). All cell lines were grown at 37 °C, 5% CO_2. All cell lines routinely tested negative for mycoplasma using a Universal Mycoplasma Detection Kit (American Type Culture Collection #30-1012K) or a MycoAlert Mycoplasma Detection Kit (Lonza #LT07-418) with the MycoAlert Assay Control Set (Lonza #LT07-518) per manufacturers' protocols.

RPE1-hTERT Cas9 *TP53*$^{-/-}$ *PTPA*$^{-/-}$ and RPE1-hTERT Cas9 *TP53*$^{-/-}$ *FANCD2*$^{-/-}$ *PTPA*$^{-/-}$ cell lines were generated using the CRISPR-Cas9 system. Oligonucleotides encoding guide RNAs targeting PTPA (5'-ACAGCAGAGGAAAGCAGCG) and a negative control (5'-GGGGCCACTAGGGACAGGAT) were cloned into the lentiCRISPRv2 vector, (Addgene #52961). Lentivirus was produced by seeding HEK293T cells at 60% confluence in 6-well plates 24 hours before transfection. The gRNA-lentiCRISPRv2 construct was transfected together with the packaging plasmids (pol/rev, VSVg, tat) using X-tremeGENE 9 DNA transfection reagent (Millipore Sigma #6365779001). Virus was harvested 48 hours after transfection and filtered through a 0.45 μm membrane. Virus was transduced into *FANCD2*$^{+/+}$ or *FANCD2*$^{-/-}$ cells with 8 μg/mL polybrene (Millipore Sigma #tr-1003-g) at a density of 10%. After 48 hours, puromycin selection was performed with 25 μg/mL puromycin (Invivogen #ant-pr-1) for 72 hours. Knockouts were validated by immunoblotting.

3.5.3 CRISPR screens

CRISPR screens were performed as previously described[76], with some modifications. RPE1-hTERT Cas9 *TP53*$^{-/-}$ and RPE1-hTERT Cas9 *TP53*$^{-/-}$ *FANCD2*$^{-/-}$ cells were transduced with the lentiviral TKOv3 library (Addgene #125517) at a low MOI (0.3), and puromycin-conditioned media was added 24 hours following transduction. The following

day, cells were trypsinized and replated in the same plates while maintaining the puromycin selection. The puromycin selection was maintained for an additional 48 hours following replating. Four days after transduction, which was considered the initial time point (t0), cells were pooled and divided into two technical replicates. Screens were performed by subculturing cells at days 3 and 6 (t3 and t6), at which point each replicate was divided into untreated and treated conditions. Cells were subcultured in medium without drug every three days (t9, t12, and t15). UVA and psoralen/UVA (PUVA) treatments were applied one day after subculturing cells (t7, t10, t13 and t16). PUVA treatment was applied according to previous determination of LD20 concentrations in uninfected cells. For PUVA treatment, plates were washed with PBS and the media replaced with Hanks' Balanced Salt Solution (HBSS), calcium, magnesium (Gibco #24020117) either containing DMSO or trimethylpsoralen (Millipore Sigma #6137) and incubated in the dark, under normal atmospheric conditions for 45 minutes. After incubation, plates were irradiated with UVA 4 X 200,000 μJ, rotating the plate 45 degrees between every dose for even coverage. After incubation, conditioned HBSS was aspirated and replaced with fresh DMEM. On t18, cell pellets were frozen for genomic DNA extraction. Screens were performed in technical duplicates and library coverage of >400 sgRNA was maintained at every passage. gDNA from cell pellets was extracted using the QIAamp Blood Maxi Kit (QIAGEN # 51194), and genome-integrated sgRNA sequences were amplified via PCR using NEBNext Ultra II Q5 Master Mix (New England Biolabs [NEB] #M5044) using primers forward: 5'-GAGGGCCTATTCCCATGATTC and reverse: 5'-CAAACCCAGGGCTGCCTTGGAA[76]. A second PCR reaction containing i5 and i7 multiplexing barcodes was completed, and products were gel purified and evaluated with

an Agilent BioAnalyzer using the High Sensitivity Kit (Agilent #5067-4626) before sequencing on Illumina NextSeq2000 to determine sgRNA representation in each sample.

3.5.4 Quantification, statistical analysis, and data quality control

The sgRNA read count files were computed from the raw CRISPR fastq files using the 'count function from the MAGeCK software[77]. The gene scores (normZ values) were estimated from the read count files using the DrugZ algorithm, testing the PUVA treated replicates against their respective non-treated controls. The MAGeCK count summary provides a summary to assess screen data quality. Both screens had a mapped reads percentage above 86% and Gini indexes below 0.095.

3.5.5 Cell viability assay

FANCD2$^{+/+}$ *PTPA*$^{+/+}$, *FANCD2*$^{+/+}$ *PTPA*$^{-/-}$, *FANCD2*$^{-/-}$ *PTPA*$^{+/+}$, and *FANCD2*$^{-/-}$ *PTPA*$^{-/-}$ cells were plated at 2000 cells per well in a 96-well plate in triplicate. 2 mM cisplatin stock solution was made by dissolving cis-Diammineplatinum(ii) dichloride (Sigma #P4394) in 154 mM NaCl, aliquoted, and frozen at -80 °C. The stock solution was diluted and titrated down to treat the cells at final cisplatin concentrations of 0, 0.04, 0.08, 0.16, 0.32, 0.63. 1.25, 2.5, 5, and 10 μM. Cells were cultured in treatment for 5 days, after which cell viablity was measured using CellTiter-Glo 2.0 (Promega #G9241).

3.5.6 Protein expression and purification

Recombinant human PP2A-AC was expressed and purified essentially as previously described[78]. The Bac-to-Bac Baculovirus Expression System (Thermo Fisher Scientific #10359016) was used to generate baculoviruses expressing 6xHis-PP2A-AC according to the manufacturer's protocols. Hi5 insect cells were transfected with the baculoviruses, and

recombinant protein was expressed in 2 L of Hi5 suspension culture. Cells were collected once the cell viability dropped to about 80%, generally after 48-72 hours. Harvested cells were lysed by sonication in lysis buffer (50 mM Tris-HCl [pH 8.0], 100 mM NaCl, 20 mM imidazole, 284 µg/mL leupeptin, 1.37 µg/mL pepstatin A, 0.17 mg/mL phenylmethylsulfonyl fluoride [PMSF], 0.33 mg/mL benzamidine, 1X Roche cOmplete EDTA-free Protease Inhibitor Cocktail). The lysate was spun at 68,900 x g in an A27-8x50 fixed angle rotor using a Sorvall Lynx 4000 centrifuge, and the soluble fraction was incubated with Ni-NTA resin (Qiagen #30230) overnight at 4 °C. After binding, the resin was washed extensively with wash buffer (50 mM Tris-HCl [pH 8.0], 500 mM NaCl, 50 mM imidazole, 10 mM ATP, 20 mM $MgCl_2$, 284 µg/mL leupeptin, 1.37 µg/mL pepstatin A, 0.17 mg/mL PMSF, 0.33 mg/mL benzamidine). His-tagged hPP2A-AC was eluted with elution buffer (50 mM Tris-HCl [pH 8.0], 300 mM NaCl, 200 mM imidazole, 5 mM β-mercaptoethanol [BME], 284 µg/mL leupeptin, 1.37 µg/mL pepstatin A, 0.17 mg/mL PMSF, 0.33 mg/mL benzamidine). The protein was buffer exchanged into storage buffer (50 mM Tris-HCl [pH 8.0], 300 mM NaCl, 1 mM DTT, 10% glycerol), aliquoted, snap frozen, and stored at -80 °C.

3.5.7 Phosphatase activity assay

Serine/threonine phosphatase activity of recombinant 6xHis-hPP2A-AC was determined using a nonradioactive molybdate dye-based phosphatase assay kit (Promega #V2460) according to the manufacturer's instructions. 5 µM of protein in assay buffer (250 mM imidazole [pH 7.2], 1 mM ethylenediaminetetraacetic acid [EDTA], 0.1% BME and 0.5 mg/mL bovine serum albumin [BSA]) was incubated with 10 µM okadaic acid or 100 µM cytostatin for 15 mins at room temperature followed by addition of 100 µM substrate

phosphothreonine peptide (RRA[pT]VA). The 50 µL reaction was set up in a 96-well plate and incubated at 37 °C for 30 mins. The reaction was stopped by adding 50 µL molybdate dye additive mix, and color was developed by incubating at room temperature for 30 mins. A standard curve was obtained using KH_2PO_4 at concentrations ranging from 2 to 40 µM. Absorbance at 600 nm was measured, and the amount of free phosphate released was quantified.

3.5.8 Preparation of Xenopus egg extracts

Animal work was performed at Caltech, which has an approved Animal Welfare Assurance (no. D16-00266) from the NIH Office of Laboratory Animal Welfare. All procedures involving *Xenopus laevis* were approved by IACUC (Protocol IA20-1797, approved May 28, 2020). Preparation of high-speed supernatant (HSS) and nucleoplasmic extract (NPE) from *Xenopus laevis* oocytes was done essentially as described previously[79]. In brief, six laboratory-bred wild-type adult female *X. laevis* (aged >2 years) were primed with 40 IU of human chorionic gonadotropin (hCG), followed by an injection of 500 IU hCG 2-8 days later. Eggs were collected the next day, dejellied in 1 L 2.2% (w/v) cysteine, pH 7.7, and washed first with 2 L 0.5X Marc's Modified Ringer's (MMR) solution (2.5 mM HEPES-KOH [pH 7.8], 50 mM NaCl, 1 mM KCl, 0.25 mM $MgSO_4$, 1.25 mM $CaCl_2$, 0.05 mM EDTA) and then with 1 L Egg Lysis Buffer (ELB) (10 mM HEPES-KOH [pH 7.7], 50 mM KCl, 2.5 mM $MgCl_2$, 250 mM sucrose, 1 mM dithiothreitol [DTT], and 50 µg/mL cycloheximide). Eggs were transferred to 14 mL round bottom Falcon tubes (Corning #352059) and packed by spinning at 200 x g for 1 minute using a Sorvall ST 8 centrifuge (Thermo Fisher Scientific). Apoptotic eggs were aspirated from each tube, and the eggs were supplemented with 5 µg/mL aprotinin, 5 µg/mL leupeptin, and 2.5 µg/mL

cytochalasin B before being crushed by centrifugation at 20,000 x g for 20 mins at 4 °C in a TH13-6x50 swinging bucket rotor using a Sorvall Lynx 4000 centrifuge. Low-speed supernatant (LSS) was collected, supplemented with 50 µg/mL cycloheximide, 1 mM DTT, 10 µg/mL aprotinin, 10 µg/mL leupeptin, and 5 µg/mL cytochalasin B, and spun at 260,000 x g for 90 mins at 2 °C in thinwall ultra-clear centrifuge tubes (Beckman Coulter #347356) in a TLS-55 swinging bucket rotor using an Optima MAX-E tabletop ultracentrifuge (Beckman Coulter). The lipids that accumulated at the top of each tube were aspirated, and the cytosolic layer underneath was recovered, aliquoted, snap frozen in liquid nitrogen, and stored at -80 °C.

NPE preparation was begun the same way, except that twenty frogs were used rather than six, and the volumes of 2.2% (w/v) cysteine, MMR, and ELB used to dejelly and wash the eggs were doubled. The protocols diverged after LSS collection – for an NPE preparation, the LSS was supplemented with 50 µg/mL cycloheximide, 1 mM DTT, 10 µg/mL aprotinin, 10 µg/mL leupeptin, 5 µg/mL cytochalasin B, and 3.3 µg/mL nocodazole. The LSS was then transferred to 14 mL Falcon tubes (Corning #352059) and spun at 20,000 x g for 15 mins at 4 °C in a TH13-6x50 swinging bucket rotor using a Sorvall Lynx 4000 centrifuge to remove precipitates. The supernatant was transferred to a 50 mL conical tube and allowed to warm to room temperature, at which point it was supplemented with 2 mM ATP, 20 mM phosphocreatine, and 5 µg/mL creatine phosphokinase. Demembranated sperm chromatin was gradually added to a final concentration of 6,600 U/µL to initiate nuclei formation, which was allowed to proceed for 75-90 mins before nuclei were separated by spinning the extract at 18,000 x g for 3 mins at 4 °C in glass culture tubes (VWR #47729-572) in a TH13-6x50 swinging bucket rotor using a Sorvall Lynx 4000 centrifuge. The

nuclei that accumulated at the top of each tube were harvested, transferred to either thickwall (Beckman Coulter #347778) or thinwall (Beckman Coulter #347356) ultracentrifuge tubes (depending on the volume of nuclei collected), and spun at 260,000 x g for 30 mins at 2 °C in a TLS-55 swinging bucket rotor using an Optima MAX-E tabletop ultracentrifuge (Beckman Coulter). The lipids that accumulated at the top of each tube were aspirated, and the nucleoplasmic layer underneath was recovered, aliquoted, snap frozen in liquid nitrogen, and stored at -80 °C.

3.5.9 Immunodepletions

Immunodepletions using antibodies raised against FANCD2 were performed essentially as described previously[37]. Briefly, Protein A Sepharose Fast Flow (Cytiva #17127903) resin was washed four times with 1X phosphate-buffered saline (PBS) by spinning at 622 x g for 30 seconds in an S-24-11-AT rotor using an Eppendorf 5430R centrifuge. These spin conditions throughout this depletion protocol, unless stated otherwise. After the washes, the resin incubated with 3.6 volumes of anti-FANCD2 serum overnight at 4 °C on a rotating wheel. The next day, the resin was washed twice with 1X PBS, once with ELB mix (2.5 mM $MgCl_2$, 50 mM KCl, 10 mM HEPES, 250 mM sucrose), twice with ELB mix supplemented with 0.5 M NaCl, and three times with ELB mix. Three rounds of depletion were performed by adding one volume of HSS or NPE to 0.14 volumes of resin and incubating on a rotating wheel for 20 mins per round at room temperature. After the final round, extract supernatants were collected and used for replication reactions, as described below.

3.5.10 Preparation of oligonucleotide duplexes with site-specific interstrand cross-links

Cross-linked oligonucleotides were prepared essentially as previously described[5,7,80]. To generate an oligonucleotide duplex containing a cisplatin-ICL, 1 mM cisplatin was incubated in 10 mM $NaClO_4$ (pH 5.2), 0.95 mM $AgNO_3$ for 24 hours at 37 °C in the dark to produce monoaquamonochloro cisplatin. Then, the cisplatin monoadduct was produced by incubating 0.125 mM Pt-ICL top oligonucleotide with 0.375 mM monoaquamonochloro cisplatin in 5.63 mM $NaClO_4$ (pH 5.2) for 12 mins at 37 °C, after which NaCl was added to 0.1 M to quench the reaction. The top oligonucleotide containing a cisplatin monoadduct was purified from a 20% acrylamide/bis (19:1), 1X TBE, 8 M urea gel by excising the gel slice containing the monoadducted oligonucleotide, crushing it, and eluting into Tris-EDTA (TE) buffer. 1.05 molar equivalents of Pt-ICL bottom oligonucleotide were added to the purified monoadducted top oligonucleotides in 100 mM $NaClO_4$, and cross-linking was promoted by incubation at 37 °C for 48 hours. The cross-linked duplex was gel-purified in the same way as the monoadduct. The purified cross-linked oligonucleotide duplex was eluted into 10 mM Tris-HCl (pH 7.4), 10 mM $NaClO_4$, snap frozen in liquid nitrogen, and stored at -80 °C.

To generate the oligonucleotide duplex containing an ICL formed through an abasic (AP) site, the corresponding top and bottom oligonucleotides (AP-ICL top and AP-ICL bottom) were annealed at a concentration of 5 μM each in 30 mM HEPES-KOH (pH 7.0), 100 mM NaCl by heating to 95 °C for 5 mins and cooling at 1 °C/min to 18 °C on a Bio-Rad T100 Thermal Cycler. The annealed duplex was then digested with 0.05 U/μL uracil-DNA glycosylase (NEB #M0280) in 1X UDG buffer (20 mM Tris-HCl, 10 mM DTT, 10 mM EDTA [pH 8.0]) for 120 mins at 37 °C, extracted with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0), precipitated in ethanol, and resuspended in 50 mM HEPES-KOH (pH

7.0), 100 mM NaCl. The cross-link was formed by incubating the oligonucleotide duplex at 37 °C for 5-7 days, after which the DNA was precipitated in ethanol, suspended in formamide loading dye (86% formamide, 1X TBE, 20 mM EDTA, 0.025% [w/v] xylene cyanol, and 0.025% [w/v] bromophenol blue), and purified from a 20% acrylamide/bis (19:1), 1X TBE, 8 M urea gel as described above. The cross-linked duplex was eluted into TE buffer, extracted with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0), precipitated in ethanol, resuspended in 10 mM Tris-HCl (pH 8.5), aliquoted, snap frozen in liquid nitrogen, and stored at -80 °C.

3.5.11 Preparation of interstrand cross-link containing plasmids (pICL)

Plasmids containing site-specific ICLs were prepared exactly as described previously[81]. $pICL^{Pt}$ and $pICL^{AP}$ were prepared by incubating 0.5 nM linear BbsI-digested plasmid with 0.75 nM Pt-ICL or AP-ICL duplex and 0.4 U/µL NEB T4 DNA ligase in 1X ligase buffer (50 mM Tris-HCl [pH 7.5], 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT) at 4 °C for 16 h. The ligation reaction was then concentrated using a 15 mL Amicon filter unit (30 kDa MWCO), extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0) and once with chloroform, ethanol precipitated, and resuspended in TE. For every 1 mL of the concentrated DNA mixture, 1 g CsCl was added to the mixture. The DNA was transferred to a 5.1 mL Quick-Seal tube (Beckman Coulter) and supplemented with 50 µg/mL ethidium bromide. The tube was filled to the top with a CsCl-TE solution consisting of 1 g of CsCl per 1 mL of TE. The mixture was spun for 23 h at 20 °C in an NVT-90 rotor at 389,340 × g. The covalently closed circular plasmid was collected and extracted six times with an equal volume of saturated n-butanol to remove ethidium bromide. The DNA was then buffer exchanged into 10 mM Tris-HCl (pH 7.5) for $pICL^{AP}$ and 10 mM Tris-HCl (pH 7.4),

10 mM $NaClO_4$ for $pICL^{Pt}$ with a 4 mL Amicon filter unit (30 kDa MWCO). Aliquots of both plasmids were frozen in liquid nitrogen and stored at −80 °C.

3.5.12 Replication reactions

Plasmids were replicated in *Xenopus* egg extracts at a final concentration of 2.5 ng/μL (except where indicated otherwise) essentially as described previously[82]. Where indicated, licensing mix was supplemented with 167-333 nM 3000 Ci/mmol [α-^{32}P]dCTP (Revvity #BLU513H250UC). Where indicated, NPE mix was supplemented with 50 μM PHA-767491, 300 μM NMS-873 (Millipore Sigma #SML1128), 7.5 mM caffeine (Millipore Sigma # C0750), 7.5-30 μM flavopiridol (MedChemExpress #HY-10005), 4 μM AZD6738 (MedChemExpress #HY-19323), or 75 μM BRC4 (Biosynth #CRB1001302). Replication was initiated by mixing 1 volume of licensing mix with 2 volumes of NPE mix. Where indicated, replication reactions were supplemented with 1-6 μM okadaic acid (Tocris Bioscience #1136) or 250 μM cytostatin (Santa Cruz Biotechnology #sc-394496) 5 minutes into the reaction; 300 μM KU-55933 (MedChemExpress #HY-12016) or 100 μM AZD0156 (MedChemExpress #HY-100016) 20 minutes into the reaction; or 0.35-35 ng/μL linear plasmid, 4 μM hPP2A-AC, 5 mM C5 (MedChemExpress #HY-128729), 3 mM C200, 300 μM F684, 20 μM flavopiridol, or 1 mM HRO761 (Selleckchem #E1908) 30 minutes into the reaction. For native agarose gel analysis, 1 μL samples of each reaction were quenched into 6 μL replication stop mix (80 mM Tris-HCl [pH 8.0], 8 mM EDTA, 0.13% phosphoric acid, 10% ficoll, 5% sodium dodecyl sulfate [SDS], 0.1% bromophenol blue) at the indicated time points. Samples were worked up and visualized by phosphorimaging as described previously[81]. For Southern blots and nascent strand analysis, 4 μL samples of each reaction were quenched into 40 μL clear stop mix (50 mM Tris-HCl

[pH 8.5], 0.5% SDS, 25 mM EDTA) at the indicated time points. Once all time points were quenched, samples were digested with 200 µg/mL RNAse A for 30 mins at 37 °C and with 1 mg/mL proteinase K for 60 mins at 37 °C. Samples were adjusted to 200 µL with 10 mM Tris-HCl pH 8.5, extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0) and once with chloroform, ethanol precipitated, and resuspended in 8-10 µL 10 mM Tris-HCl pH 8.5. Samples were subsequently digested and analyzed as described below for their respective assays.

3.5.13 Southern blotting

Southern blotting was performed exactly as described previously[81]. Samples were replicated as described above, and 3 µL of recovered DNA were incubated with 5 U HincII in 1X NEBuffer rCutsmart (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 µg/mL recombinant albumin; pH 7.9) in a total volume of 5 µL at 37 °C for 60 min. Following digestion, 5 µL of 50 mM Tris-HCl (pH 8.8), 50 mM NaCl, 10 mM EDTA, and 2 µL of 6X alkaline loading buffer (300 mM NaOH, 6 mM EDTA, 18% ficoll, 0.25% xylene cyanol, 0.15% bromocresol green) was added to each sample. The DNA was resolved on a denaturing 0.8% agarose, 50 mM NaOH, 1 mM EDTA gel for 20 h at 1.25 V cm^{-1}. After running was complete, the gel was incubated in 0.25 M HCl for 20 min and in 0.4 M NaOH, 3 M NaCl for 60 min. Downward capillary transfer onto an Invitrogen BrightStar-Plus positively charged nylon membrane was performed overnight in 0.4 M NaOH, 1.5 M NaCl. The membrane was subsequently washed with 4X SSC buffer (0.6 M NaCl, 60 mM sodium citrate; pH 7.0) for 5 min, and the DNA was cross-linked to the membrane via 265 nm UV irradiation (120,000 µJ/cm^2) with a UVP HybriLinker Oven. The membrane was then incubated with 15 mL ULTRAhyb Ultrasensitive Hybridization

Buffer (Invitrogen) for 1-2 h at 42 °C. During this prehybridization, a radioactive probe was made using 50 ng of linearized pCtrl and the Prime-a-Gene Labeling System kit (Promega) following the manufacturer's instructions. The probe was added to the hybridization tube and incubated with the membrane at 42 °C for 18 h. After hybridization, the membrane was washed twice with 1X SSC, 0.7% SDS at 42 °C for 10-20 min per wash and twice with 0.1X SSC, 0.7% SDS that was pre-heated to 75 °C for 10-20 min per wash. The membrane was dried briefly on blotting paper and visualized by phosphorimaging.

3.5.14 Nascent strand analysis

Nascent strand analysis was performed essentially as described previously[12]. Briefly, plasmids were replicated and purified as described above. For leading strand analysis, 2 µL of recovered DNA was digested with 2 U AflIII (NEB #R0541) and 2 U of either BamHI (NEB #R0136) or EcoRI-HF (NEB #R3101) at 37 °C for 2.5-4 hours. For lagging strand analysis[83], 2 µL DNA was digested with 2 U Nt.BstNBI (NEB #R0607) at 37 °C for 2.5-4 hours. Denaturing formamide gel loading buffer (Invitrogen #AM8547) was added 1:1 to each reaction, and nascent strands were resolved on a 7% acrylamide/bis (19:1), 8 M urea, 0.8X GTG buffer (71 mM Tris-HCl, 23 mM taurine, 0.4 mM EDTA) gel. Gels were dried and visualized by phosphorimaging.

3.5.15 Plasmid pulldown

Plasmid pulldowns were performed essentially as previously described[11]. In brief, 4.5 µL Dynabeads Protein A beads (Invitrogen #10001D) per time point were washed 3 times with wash buffer (50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1 mM EDTA [pH 8.0], 0.02% Tween 20) on a magnetic rack and bound to 2 pmol biotinylated LacI per µL of beads for

40 mins at room temperature on a rotating wheel. Following the incubation, dynabeads were washed 4 times with pulldown buffer (2.5 mM $MgCl_2$, 50 mM KCl, 10 mM HEPES, 250 mM sucrose, 0.25 mg/mL BSA, 0.02% Tween 20) on a magnetic rack, aliquoted for individual time points, and stored on ice. Replication reactions were performed as described above, but at a plasmid concentration of 5 ng/μL in extracts. At the indicated times, 6 μL of each reaction were quenched into 30 μL Dynabeads and incubated on a rotating wheel for 30 mins at 4 °C. Following the incubation, the beads were washed 3 times with 2.5 mM $MgCl_2$, 50 mM KCl, 10 mM HEPES, 0.25 mg/mL BSA, 0.03% Tween 20 on a magnetic rack, suspended in 1X Laemmli loading buffer, and analyzed by immunoblotting. 1 μL of each replication reaction was also quenched directly into 1X Laemmli loading buffer to serve as an input sample.

3.5.16 Immunoblotting

Samples in 1X Laemmli loading buffer, generally corresponding to 0.3-0.75 μL of replication reaction (except for plasmid pulldowns, which are described above), were resolved on 4-15% (Bio-Rad #5671085, #4561086) or 10% (Bio-Rad #5671035, #4561036) acrylamide gels and transferred to 0.45 μM polyvinyl difluoride (PVDF) membrane (Thermo Fisher Scientific #88518). Membrane was blocked with 5% nonfat milk in 1X PBS with 0.05% Tween 20 (PBST) for 60 mins at room temperature, rinsed with 1X PBST, and incubated in primary antibody in 1X PBST overnight at 4 °C. Commercial primary antibody identities and dilutions used for blotting cell samples were as follows: rabbit polyclonal anti-PTPA (human), Cell Signaling Technology #3330, 1:1,000; mouse monoclonal anti-FANCD2 (human), Santa Cruz Biotechnology #sc-20022, 1:1,000; and mouse monoclonal anti-RPA (human), Millipore Sigma #MABE285, 1:1,000.

Commercial primary antibody identities and dilutions used for blotting *Xenopus* egg extract samples were as follows: rabbit monoclonal anti-phospho-ATM (S1981; human), Rockland Immunochemicals #200-301-500, 1:1,500; rabbit polyclonal anti-H3 (human), Cell Signaling Technology #9715, 1:500; rabbit polyclonal anti-phospho-CHK1 (S345; human), Cell Signaling Technology #2341, 1:500; rabbit polyclonal anti-phospho-RPA32 (S33; human), Bethyl Laboratories #A300-246A, 1:1,000; and mouse monoclonal anti-RAD51, Novus Biologicals #NB100-148, 1:1,000. Rabbit polyclonal anti-CHK2[84], rabbit polyclonal anti-FANCD2[12], rabbit polyclonal anti-RPA70[85], and rabbit polyclonal anti-MRE11[86] have been previously described. Rabbit polyclonal anti-DNA2 and rabbit polyclonal anti-WRN were raised against fragments corresponding to amino acids 1-259 and 242-472 of the *X. laevis* proteins, respectively. Anti-DNA2 was diluted 1:1,000, and anti-WRN was diluted 1:5,000. The next day, membrane was washed 3 times with 1X PBST and incubated in secondary antibody diluted 1:25,000 in 5% nonfat milk in 1X PBST for 30 mins at room temperature. Secondary antibody was either goat polyclonal anti-rabbit IgG (H+L) conjugated to horseradish peroxidase (HRP), Jackson ImmunoResearch #111-035-003 or goat polyclonal anti-mouse IgG (H+L) conjugated to HRP, Jackson ImmunoResearch #115-035-003. Following the secondary antibody incubation, membranes were washed 3 times with 1X PBST, incubated for 2-4 mins in chemiluminescence substrate (either Thermo Fisher Scientific #34076 or Genesee Scientific #20-300S), and imaged using a ChemiDoc Imaging System (Bio-Rad).

3.5.17 Sample preparation for phosphoproteomics

Plasmids were replicated at a final concentration of 5 ng/µL in *Xenopus* egg extract in the absence or presence of okadaic acid as described above. After 90 minutes of replication,

40 µL samples of each reaction were quenched in triplicate into 400 µL SDS solubilization buffer (5% SDS, 50 mM triethylammonium bicarbonate [TEAB; pH 8.5]) each. Proteins were reduced by adding tris(2-carboxyethyl)phosphine (TCEP) to 5 mM and incubating at 55 °C for 20 mins. After samples cooled to room temperature, proteins were alkylated by adding chloroacetaldehyde (CAA) to 100 mM and incubating in the dark for 30 mins. Samples were spun at 13,000 x g for 8 mins in an Eppendorf 5424R microcentrifuge to remove undissolved matter, and trifluoroacetic acid (TFA) was added to 1%. Sample volume was increased approximately 7-fold by adding S-Trap buffer (90% methanol, 100 mM TEAB [pH 7.55]), and samples were loaded onto S-Trap Midi Columns (ProtiFi). Protein cleanup, on-column digestion with 50 µg trypsin per column, and peptide elution was performed according to the manufacturer's instructions, and eluted peptides were dried by SpeedVac.

Peptides were reconstituted in 1 mL 1% TFA. Sep-Pak tC18 cartridges (Waters Corp #WAT054960) were placed into 15 mL conical tubes and activated by adding 1 mL 100% acetonitrile (ACN) and spinning at 50 x g for 1 min at room temperature. Note that these spin conditions were also used for all subsequent washes of the Sep-Pak columns. The ACN wash was repeated 3 additional times (4 total washes), after which the columns were equilibrated by washing 4 times with 1 mL 0.1% TFA. The acidified peptide samples were loaded onto the columns and allowed to drain by gravity, and the collected flowthrough was reapplied once. Columns were washed 4 times with 1 mL 0.1% formic acid (FA), and peptides were eluted by gravity 3 times with 350 µL 0.1% FA in 50% ACN. Samples were dried by SpeedVac, after which phosphoenrichment was performed using the Thermo Fisher Scientific TiO_2 Phosphopeptide Enrichment Kit (#A32993) according to the

manufacturer's instructions. The flowthroughs from the binding and wash steps of the TiO_2 enrichment were pooled and used for another round of phosphoenrichment with the Thermo Fisher Scientific Fe-NTA Phosphopeptide Enrichment Kit (#A32992) according to the manufacturer's instructions. Eluates from the two rounds of phosphoenrichment were separately dried by SpeedVac and reconstituted in 0.1% FA.

LC-MS/MS data were acquired in data-independent acquisition (DIA) mode using a Thermo Fisher Scientific Vanquish Neo Tandem UHPLC system coupled to an Orbitrap Astral mass spectrometer. A total of 0.5 µg peptides was injected onto a PepMap Neo Trap Cartridge (5 cm x 300 µm, 5 µm C18, Thermo Fisher Scientific) before refocusing on an Aurora UHPLC column (25 cm x 75 µm, 1.7 µm C18, AUR3-25075C18-XT, Ion Opticks) with a flow rate of 0.25 µL/min in a 35-min gradient. The gradient was composed of 5-10% B (1 min, 0.5 µL/min), 10% B (0.1 min, 0.25 µL/min), 10-30% B (25.5 min, 0.25 µL/min), 30-44% B (3 min, 0.25 µL/min), 44-99% B (2 min, 0.35 µL/min), and 99% B (3.4 min, 0.5 µL/min); mobile phase A: 0.1% FA in water; mobile phase B: 80% ACN, 0.1% FA. The Orbitrap Astral MS was operated at a full MS resolution of 240,000 with a full scan range of 380-980 m/z. The full MS AGC was set to 300%, and IT was set to 10 ms. MS/MS scans were recorded with a 2 Th isolation window, 3.5 ms maximum ion injection time. The normalized AGC target was set to 500%, and the isolated ions were fragmented using HCD with 27% NCE.

3.5.18 Proteomic data analysis

RAW files were processed using the CHIMERYS[87] intelligent search algorithm with Inferys 4.7.0 fragmentation prediction (version 4.0, MSAID GmbH) in Proteome

Discoverer (PD) (v3.2; Thermo Fisher Scientific) against the 10.1 version of the *Xenopus laevis* genome downloaded from Xenbase:

https://wuehr.scholar.princeton.edu/sites/g/files/toruqf2761/files/xenla_10_1_xenbase.fasta.txt. The maximum missed cleavages was set to 2. Dynamic modifications were set to oxidation on methionine (M) and phosphorylation on serine, threonine, and tyrosine (S, T and Y). Carbamidomethylation on cysteine residues (C) was set as a fixed modification. Peptides between 7 and 30 amino acids were considered, and the false discovery rate (FDR) for peptide spectra matches (PSMs) identifications was set to 1%. Quantification type was set to MS2 Apex. Data were deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD071242[88]. Downstream analysis was performed using the R packages MSstats (v4.16.1)[89] and MSstatsPTM (v2.10.3)[90]. Peptide-spectrum match (PSM) tables were imported and processed using the PDtoMSstatsPTMFormat() function. A custom function was implemented to parse Xenbase FASTA headers to extract protein accession identifiers compatible with MSstatsPTM, and peptide abundances were Log_2-transformed and summarized at the phosphosite and protein levels. Differential abundance analysis between control and okadaic acid-treated samples was performed using linear mixed-effects models, which estimated Log_2 fold changes (Log_2FCs) and p-values for each phosphosite. Resulting p-values were adjusted for multiple hypothesis testing using the Benjamini-Hochberg procedure[91]. Phosphosites with FDR-adjusted p-values < 0.05 and absolute $Log_2FC > 0.5$ were considered significantly regulated.

3.5.19 Preparation and analysis of next-generation sequencing libraries

Next-generation sequencing (NGS) libraries were prepared and analyzed exactly as described previously[81]. Raw next-generation sequencing reads were deposited to the NCBI Sequence Read Archive with the BioProject ID PRJNA1367151. Raw sequencing reads were trimmed to 150 bp with Trimmomatic (v0.39)[92], processed with SIQ (v1.8) using default settings with long-read analysis enabled[93], and visualized with SIQPlotteR.

3.6 Acknowledgements

We thank members of the Semlow lab for helpful discussions and comments on the manuscript. We thank D. Durocher for providing the RPE1-hTERT Cas9 *TP53*$^{-/-}$ cell line used for chemogenomic screening. We also thank I. Fianu for the PP2A expression construct, B. Shen for the EXO1 inhibitors, W. Dunphy for the RPA70 antibody, and J. Walter for the MRE11 and FANCD2 antibodies. D.R.S. is supported by NIH grant no. R01GM151410 and a Shurl and Kay Curci Foundation Research Grant and is a Ronald and JoAnne Willens Scholar and a Vallee Scholar.

3.7 Author contributions

M.V.A. performed the experiments described in Figs. 3.2a-e,g, 3.3, 3.4, and 3.5 and Supplementary Figs. 3.2a,b,f, 3.3e-g, 3.4, and 3.5. V.A.M. performed the experiments described in Figs. 3.1b,c and 3.2f and Supplementary Figs. 3.1, 3.2e, and 3.3a,b. J.T. performed the experiments described in Fig. 3.1d,e. R.N. purified and determined the activity of the recombinant hPP2A-AC complex, as described in Supplementary Fig. 3.2c,d. J.K.K. performed the experiments described in Supplementary Fig. 3.3c,d. T.Y.W.,T.F.C., and V.A.M. assisted with performing and analyzing the mass spectrometry data described in Fig. 3.5a. E.J.G., O.E.O., and J.M.D. prepared DNA2 and WRN

antibodies. M.V.A., V.A.M., and D.R.S. designed and interpreted the experiments with input from J.L.C. M.V.A. and D.R.S. wrote the manuscript with input from all authors.

3.8 Declaration of interests

The authors declare no competing interests.

3.9 References

1 Semlow, D. R. & Walter, J. C. Mechanisms of Vertebrate DNA Interstrand Cross-Link Repair. *Annu Rev Biochem* **90**, 107-135 (2021). https://doi.org/10.1146/annurev-biochem-080320-112510

2 Deans, A. J. & West, S. C. DNA interstrand crosslink repair and cancer. *Nat Rev Cancer* **11**, 467-480 (2011). https://doi.org/10.1038/nrc3088

3 Thongthip, S. *et al.* Fan1 deficiency results in DNA interstrand cross-link repair defects, enhanced tissue karyomegaly, and organ dysfunction. *Genes Dev* **30**, 645-659 (2016). https://doi.org/10.1101/gad.276261.115

4 Magaña-Schwencke, N., Henriques, J. A., Chanet, R. & Moustacchi, E. The fate of 8-methoxypsoralen photoinduced crosslinks in nuclear and mitochondrial yeast DNA: comparison of wild-type and repair-deficient strains. *Proc Natl Acad Sci U S A* **79**, 1722-1726 (1982). https://doi.org/10.1073/pnas.79.6.1722

5 Zhang, J. *et al.* DNA interstrand cross-link repair requires replication-fork convergence. *Nat Struct Mol Biol* **22**, 242-247 (2015). https://doi.org/10.1038/nsmb.2956

6 Fullbright, G., Rycenga, H. B., Gruber, J. D. & Long, D. T. p97 Promotes a Conserved Mechanism of Helicase Unloading during DNA Cross-Link Repair. *Mol Cell Biol* **36**, 2983-2994 (2016). https://doi.org/10.1128/MCB.00434-16

7 Semlow, D. R., Zhang, J., Budzowska, M., Drohat, A. C. & Walter, J. C. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase. *Cell* **167**, 498-511.e414 (2016). https://doi.org/10.1016/j.cell.2016.09.008

8 Wu, R. A. *et al.* TRAIP is a master regulator of DNA interstrand crosslink repair. *Nature* **567**, 267-272 (2019). https://doi.org/10.1038/s41586-019-1002-0

9 Amunugama, R. *et al.* Replication Fork Reversal during DNA Interstrand Crosslink Repair Requires CMG Unloading. *Cell Rep* **23**, 3419-3428 (2018). https://doi.org/10.1016/j.celrep.2018.05.061

10 Abdullah, U. B. *et al.* RPA activates the XPF-ERCC1 endonuclease to initiate processing of DNA interstrand crosslinks. *EMBO J* **36**, 2047-2060 (2017). https://doi.org/10.15252/embj.201796664

11 Budzowska, M., Graham, T. G., Sobeck, A., Waga, S. & Walter, J. C. Regulation of the Rev1-pol ζ complex during bypass of a DNA interstrand cross-link. *EMBO J* **34**, 1971-1985 (2015). https://doi.org/10.15252/embj.201490878

12 Räschle, M. *et al.* Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* **134**, 969-980 (2008). https://doi.org/10.1016/j.cell.2008.08.030

13 Long, D. T., Räschle, M., Joukov, V. & Walter, J. C. Mechanism of RAD51-dependent DNA interstrand cross-link repair. *Science* **333**, 84-87 (2011). https://doi.org/10.1126/science.1204258

14 Gari, K., Décaillet, C., Delannoy, M., Wu, L. & Constantinou, A. Remodeling of DNA replication structures by the branch point translocase FANCM. *Proc Natl Acad Sci U S A* **105**, 16107-16112 (2008). https://doi.org/10.1073/pnas.0804777105

15 Bugreev, D. V., Rossi, M. J. & Mazin, A. V. Cooperation of RAD51 and RAD54 in regression of a model replication fork. *Nucleic Acids Res* **39**, 2153-2164 (2011). https://doi.org/10.1093/nar/gkq1139

16 Machwe, A., Karale, R., Xu, X., Liu, Y. & Orren, D. K. The Werner and Bloom syndrome proteins help resolve replication blockage by converting (regressed) holliday junctions to functional replication forks. *Biochemistry* **50**, 6774-6788 (2011). https://doi.org/10.1021/bi2001054

17 Berti, M. *et al.* Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition. *Nat Struct Mol Biol* **20**, 347-354 (2013). https://doi.org/10.1038/nsmb.2501

18 Thangavel, S. *et al.* DNA2 drives processing and restart of reversed replication forks in human cells. *J Cell Biol* **208**, 545-562 (2015). https://doi.org/10.1083/jcb.201406100

19 Symington, L. S. & Gautier, J. Double-strand break end resection and repair pathway choice. *Annu Rev Genet* **45**, 247-271 (2011). https://doi.org/10.1146/annurev-genet-110410-132435

20 Matsuoka, S. *et al.* ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160-1166 (2007). https://doi.org/10.1126/science.1140321

21 Blackford, A. N. & Jackson, S. P. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol Cell* **66**, 801-817 (2017). https://doi.org/10.1016/j.molcel.2017.05.015

22 Lee, J. H. & Paull, T. T. ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science* **308**, 551-554 (2005). https://doi.org/10.1126/science.1108297

23 Cejka, P. & Symington, L. S. DNA End Resection: Mechanism and Control. *Annu Rev Genet* **55**, 285-307 (2021). https://doi.org/10.1146/annurev-genet-071719-020312

24 Ahlskog, J. K., Larsen, B. D., Achanta, K. & Sorensen, C. S. ATM/ATR-mediated phosphorylation of PALB2 promotes RAD51 function. *EMBO Rep* **17**, 671-681 (2016). https://doi.org/10.15252/embr.201541455

25 Zhang, H. *et al.* ATM priming and end resection-coupled phosphorylation of MRE11 is important for fork protection and replication restart. *Proc Natl Acad Sci U S A* **122**, e2422720122 (2025). https://doi.org/10.1073/pnas.2422720122

26 Fugger, K. *et al.* FBH1 Catalyzes Regression of Stalled Replication Forks. *Cell Rep* **10**, 1749-1757 (2015). https://doi.org/10.1016/j.celrep.2015.02.028

27 Somyajit, K. *et al.* Redox-sensitive alteration of replisome architecture safeguards genome integrity. *Science* **358**, 797-802 (2017). https://doi.org/10.1126/science.aao3172

28 Guo, Z., Kozlov, S., Lavin, M. F., Person, M. D. & Paull, T. T. ATM activation by oxidative stress. *Science* **330**, 517-521 (2010). https://doi.org/10.1126/science.1192912

29 Olivieri, M. *et al.* A Genetic Map of the Response to DNA Damage in Human Cells. *Cell* **182**, 481-496 e421 (2020). https://doi.org/10.1016/j.cell.2020.05.040

30 Li, N. *et al.* Cooperation of the NEIL3 and Fanconi anemia/BRCA pathways in interstrand crosslink repair. *Nucleic Acids Res* **48**, 3014-3028 (2020). https://doi.org/10.1093/nar/gkaa038

31 Semlow, D. R., Zhang, J., Budzowska, M., Drohat, A. C. & Walter, J. C. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase. *Cell* **167**, 498-511 e414 (2016). https://doi.org/10.1016/j.cell.2016.09.008

32 Yoshikiyo, K. *et al.* KIAA1018/FAN1 nuclease protects cells against genomic instability induced by interstrand cross-linking agents. *Proc Natl Acad Sci U S A* **107**, 21553-21557 (2010). https://doi.org/10.1073/pnas.1011081107

33 Fellner, T. *et al.* A novel and essential mechanism determining specificity and activity of protein phosphatase 2A (PP2A) in vivo. *Genes Dev* **17**, 2138-2150 (2003). https://doi.org/10.1101/gad.259903

34 Arafeh, R., Shibue, T., Dempster, J. M., Hahn, W. C. & Vazquez, F. The present and future of the Cancer Dependency Map. *Nat Rev Cancer* **25**, 59-73 (2025). https://doi.org/10.1038/s41568-024-00763-x

35 Yang, D. *et al.* PP2A licenses the FANCD2/FANCI complex for chromosome loading. *Cell Rep* **43**, 114971 (2024). https://doi.org/10.1016/j.celrep.2024.114971

36 Chou, D. M., Petersen, P., Walter, J. C. & Walter, G. Protein phosphatase 2A regulates binding of Cdc45 to the prereplication complex. *J Biol Chem* **277**, 40520-40527 (2002). https://doi.org/10.1074/jbc.M207226200

37 Knipscheer, P. *et al.* The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* **326**, 1698-1701 (2009). https://doi.org/10.1126/science.1182372

38 Deng, L. *et al.* Mitotic CDK Promotes Replisome Disassembly, Fork Breakage, and Complex DNA Rearrangements. *Mol Cell* **73**, 915-929.e916 (2019). https://doi.org/10.1016/j.molcel.2018.12.021

39 Kawada, M., Amemiya, M., Ishizuka, M. & Takeuchi, T. Cytostatin, an inhibitor of cell adhesion to extracellular matrix, selectively inhibits protein phosphatase 2A. *Biochim Biophys Acta* **1452**, 209-217 (1999). https://doi.org/10.1016/s0167-4889(99)00126-3

40 Ambjørn, S. M. *et al.* A complex of BRCA2 and PP2A-B56 is required for DNA repair by homologous recombination. *Nat Commun* **12**, 5748 (2021). https://doi.org/10.1038/s41467-021-26079-0

41 Brewer, A. *et al.* Mapping the substrate landscape of protein phosphatase 2A catalytic subunit PPP2CA. *iScience* **27**, 109302 (2024). https://doi.org/10.1016/j.isci.2024.109302

42 Peterson, S. E. *et al.* Activation of DSB processing requires phosphorylation of CtIP by ATR. *Mol Cell* **49**, 657-667 (2013). https://doi.org/10.1016/j.molcel.2012.11.020

43 Bakkenist, C. J. & Kastan, M. B. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature* **421**, 499-506 (2003). https://doi.org/10.1038/nature01368

44 Kozlov, S. V. *et al.* Involvement of novel autophosphorylation sites in ATM activation. *EMBO J* **25**, 3504-3514 (2006). https://doi.org/10.1038/sj.emboj.7601231

45 Goodarzi, A. A. *et al.* Autophosphorylation of ataxia-telangiectasia mutated is regulated by protein phosphatase 2A. *EMBO J* **23**, 4451-4461 (2004). https://doi.org/10.1038/sj.emboj.7600455

46 Saldivar, J. C., Cortez, D. & Cimprich, K. A. The essential kinase ATR: ensuring faithful duplication of a challenging genome. *Nat Rev Mol Cell Biol* **18**, 622-636 (2017). https://doi.org/10.1038/nrm.2017.67

47 Lemacon, D. *et al.* MRE11 and EXO1 nucleases degrade reversed forks and elicit MUS81-dependent fork rescue in BRCA2-deficient cells. *Nat Commun* **8**, 860 (2017). https://doi.org/10.1038/s41467-017-01180-5

48 Liu, W. *et al.* A Selective Small Molecule DNA2 Inhibitor for Sensitization of Human Cancer Cells to Chemotherapy. *EBioMedicine* **6**, 73-86 (2016). https://doi.org/10.1016/j.ebiom.2016.02.043

49 Wang, Y. *et al.* Discovery and Characterization of Small Molecule Inhibitors Targeting Exonuclease 1 for Homologous Recombination-Deficient Cancer Therapy. *ACS Chem Biol* (2025). https://doi.org/10.1021/acschembio.5c00117

50 Ferretti, S. *et al.* Discovery of WRN inhibitor HRO761 with synthetic lethality in MSI cancers. *Nature* **629**, 443-449 (2024). https://doi.org/10.1038/s41586-024-07350-y

51 Adolph, M. B. & Cortez, D. Mechanisms and regulation of replication fork reversal. *DNA Repair (Amst)* **141**, 103731 (2024). https://doi.org/10.1016/j.dnarep.2024.103731

52 Wawrousek, K. E. *et al.* Xenopus DNA2 is a helicase/nuclease that is found in complexes with replication proteins And-1/Ctf4 and Mcm10 and DSB response proteins Nbs1 and ATM. *Cell Cycle* **9**, 1156-1166 (2010). https://doi.org/10.4161/cc.9.6.11049

53 Bolderson, E. *et al.* Phosphorylation of Exo1 modulates homologous recombination repair of DNA double-strand breaks. *Nucleic Acids Res* **38**, 1821-1831 (2010). https://doi.org/10.1093/nar/gkp1164

54 Palermo, V. *et al.* Switch-like phosphorylation of WRN integrates end-resection with RAD51 metabolism at collapsed replication forks. *Nucleic Acids Res* **52**, 12334-12350 (2024). https://doi.org/10.1093/nar/gkae807

55 Beamish, H. *et al.* Functional link between BLM defective in Bloom's syndrome and the ataxia-telangiectasia-mutated protein, ATM. *J Biol Chem* **277**, 30515-30523 (2002). https://doi.org/10.1074/jbc.M203801200

56 Sturzenegger, A. *et al.* DNA2 cooperates with the WRN and BLM RecQ helicases to mediate long-range DNA end resection in human cells. *J Biol Chem* **289**, 27314-27326 (2014). https://doi.org/10.1074/jbc.M114.578823

57 Nimonkar, A. V. *et al.* BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev* **25**, 350-362 (2011). https://doi.org/10.1101/gad.2003811

58 Higgs, M. R. *et al.* BOD1L Is Required to Suppress Deleterious Resection of Stressed Replication Forks. *Mol Cell* **59**, 462-477 (2015). https://doi.org/10.1016/j.molcel.2015.06.007

59 Ammazzalorso, F., Pirzio, L. M., Bignami, M., Franchitto, A. & Pichierri, P. ATR and ATM differently regulate WRN to prevent DSBs at stalled replication forks and promote replication fork recovery. *EMBO J* **29**, 3156-3169 (2010). https://doi.org/10.1038/emboj.2010.205

60 Wang, D. *et al.* ATM-phosphorylated SPOP contributes to 53BP1 exclusion from chromatin during DNA replication. *Sci Adv* **7** (2021). https://doi.org/10.1126/sciadv.abd9208

61 Gupta, A. *et al.* MOF phosphorylation by ATM regulates 53BP1-mediated double-strand break repair pathway choice. *Cell Rep* **8**, 177-189 (2014). https://doi.org/10.1016/j.celrep.2014.05.044

62 Wang, A. T. *et al.* A Dominant Mutation in Human RAD51 Reveals Its Function in DNA Interstrand Crosslink Repair Independent of Homologous Recombination. *Mol Cell* **59**, 478-490 (2015). https://doi.org/10.1016/j.molcel.2015.07.009

63 Schlacher, K., Wu, H. & Jasin, M. A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. *Cancer Cell* **22**, 106-116 (2012). https://doi.org/10.1016/j.ccr.2012.05.015

64 Garzon, J., Ursich, S., Lopes, M., Hiraga, S. I. & Donaldson, A. D. Human RIF1-Protein Phosphatase 1 Prevents Degradation and Breakage of Nascent DNA on

Replication Stalling. *Cell Rep* **27**, 2558-2566 e2554 (2019). https://doi.org/10.1016/j.celrep.2019.05.002

65 Mukherjee, C. *et al.* RIF1 promotes replication fork protection and efficient restart to maintain genome stability. *Nat Commun* **10**, 3287 (2019). https://doi.org/10.1038/s41467-019-11246-1

66 Cejka, P. *et al.* DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2. *Nature* **467**, 112-116 (2010). https://doi.org/10.1038/nature09355

67 Hu, J. *et al.* The intra-S phase checkpoint targets Dna2 to prevent stalled replication forks from reversing. *Cell* **149**, 1221-1232 (2012). https://doi.org/10.1016/j.cell.2012.04.030

68 Adamo, A. *et al.* Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. *Mol Cell* **39**, 25-35 (2010). https://doi.org/10.1016/j.molcel.2010.06.026

69 Pace, P. *et al.* Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. *Science* **329**, 219-223 (2010). https://doi.org/10.1126/science.1192277

70 Karanja, K. K., Cox, S. W., Duxin, J. P., Stewart, S. A. & Campbell, J. L. DNA2 and EXO1 in replication-coupled, homology-directed repair and in the interplay between HDR and the FA/BRCA network. *Cell Cycle* **11**, 3983-3996 (2012). https://doi.org/10.4161/cc.22215

71 Karanja, K. K., Lee, E. H., Hendrickson, E. A. & Campbell, J. L. Preventing over-resection by DNA2 helicase/nuclease suppresses repair defects in Fanconi anemia cells. *Cell Cycle* **13**, 1540-1550 (2014). https://doi.org/10.4161/cc.28476

72 Cai, M. Y. *et al.* Cooperation of the ATM and Fanconi Anemia/BRCA Pathways in Double-Strand Break End Resection. *Cell Rep* **30**, 2402-2415.e2405 (2020). https://doi.org/10.1016/j.celrep.2020.01.052

73 Leonard, D. *et al.* Selective PP2A Enhancement through Biased Heterotrimer Stabilization. *Cell* **181**, 688-701.e616 (2020). https://doi.org/10.1016/j.cell.2020.03.038

74 Zimmermann, M. *et al.* CRISPR screens identify genomic ribonucleotides as a source of PARP-trapping lesions. *Nature* **559**, 285-289 (2018). https://doi.org/10.1038/s41586-018-0291-z

75 Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* **42**, e168 (2014). https://doi.org/10.1093/nar/gku936

76 Olivieri, M. *et al.* A Genetic Map of the Response to DNA Damage in Human Cells. *Cell* **182**, 481-496.e421 (2020). https://doi.org/10.1016/j.cell.2020.05.040

77 Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* **15**, 554 (2014). https://doi.org/10.1186/s13059-014-0554-4

78 Fianu, I. *et al.* Structural basis of Integrator-mediated transcription regulation. *Science* **374**, 883-887 (2021). https://doi.org/10.1126/science.abk0154

79 Sparks, J. & Walter, J. C. Extracts for Analysis of DNA Replication in a Nucleus-Free System. *Cold Spring Harb Protoc* **2019** (2019). https://doi.org/10.1101/pdb.prot097154

80 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). https://doi.org/10.1007/978-1-61779-998-3_15

81 Altshuller, M. *et al.* The Fanconi anemia pathway repairs colibactin-induced DNA interstrand cross-links. *Nat Commun* **16**, 10563 (2025). https://doi.org/10.1038/s41467-025-65606-1

82 Lebofsky, R., Takahashi, T., and Walter, J.C. in *Methods in Molecular Biology, DNA Replication* Vol. 521 (ed Sonya Vengrova and Jacob Z. Dalgaard) Ch. 13, 229-252 (Humana Press, 2009).

83 Dewar, J. M., Budzowska, M. & Walter, J. C. The mechanism of DNA replication termination in vertebrates. *Nature* **525**, 345-350 (2015). https://doi.org/10.1038/nature14887

84 Guo, Z. & Dunphy, W. G. Response of Xenopus Cds1 in cell-free extracts to DNA templates with double-stranded ends. *Mol Biol Cell* **11**, 1535-1546 (2000). https://doi.org/10.1091/mbc.11.5.1535

85 Lee, J., Kumagai, A. & Dunphy, W. G. Claspin, a Chk1-regulatory protein, monitors DNA replication on chromatin independently of RPA, ATR, and Rad17. *Mol Cell* **11**, 329-340 (2003). https://doi.org/10.1016/s1097-2765(03)00045-5

86 Sathyanarayana, P., Wu, Y., Walter, J. C. & Loparo, J. J. The NHEJ machinery is translocated off DNA ends to enable resection. *bioRxiv* (2025). https://doi.org/10.1101/2025.05.13.653820

87 Frejno, M. *et al.* Unifying the analysis of bottom-up proteomics data with CHIMERYS. *Nat Methods* **22**, 1017-1027 (2025). https://doi.org/10.1038/s41592-025-02663-w

88 Perez-Riverol, Y. *et al.* The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* **50**, D543-D552 (2022). https://doi.org/10.1093/nar/gkab1038

89 Choi, M. *et al.* MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* **30**, 2524-2526 (2014). https://doi.org/10.1093/bioinformatics/btu305

90 Kohler, D. *et al.* MSstatsPTM: Statistical Relative Quantification of Posttranslational Modifications in Bottom-Up Mass Spectrometry-Based Proteomics. *Mol Cell Proteomics* **22**, 100477 (2023). https://doi.org/10.1016/j.mcpro.2022.100477

91 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B, Methodological* **57**, 289-300 (1995). https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

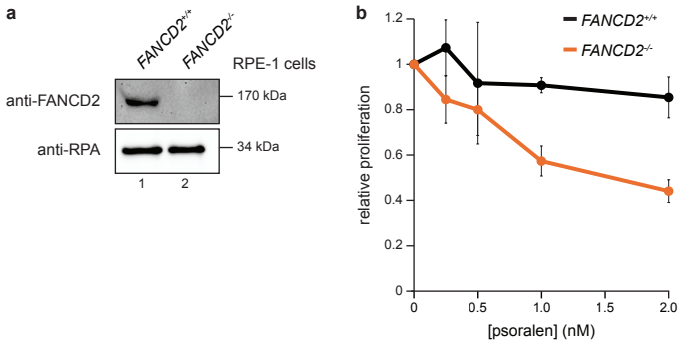
92 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014). https://doi.org/10.1093/bioinformatics/btu170

93 van Schendel, R., Schimmel, J. & Tijsterman, M. SIQ: easy quantitative measurement of mutation profiles in sequencing data. *NAR Genom Bioinform* **4**, lqac063 (2022). https://doi.org/10.1093/nargab/lqac063

3.10 Supplementary Information

Name	Sequence	Source
RA302	CTC TCC TGA CTA CTC CCA GTC A	IDT
RA303	GGC GGG ACT ATG GTT GCT	IDT
Pt-ICL top	5'-P CCC TCT TCC GCT CTT CTT TC	IDT
Pt-ICL bottom	5'-P GCA CGA AAG AAG AGC GGA AG	IDT
AP-ICL top	5'-P CCC TCT TCC GCT CdUT CTT TC	IDT
AP-ICL bottom	5'-P GCA CGA AAG AAG AGC GGA AG	IDT
seq ladder primer	CAT GTT TTA CTA GCC AGA TTT TTC CTC CTC TCC TG	IDT

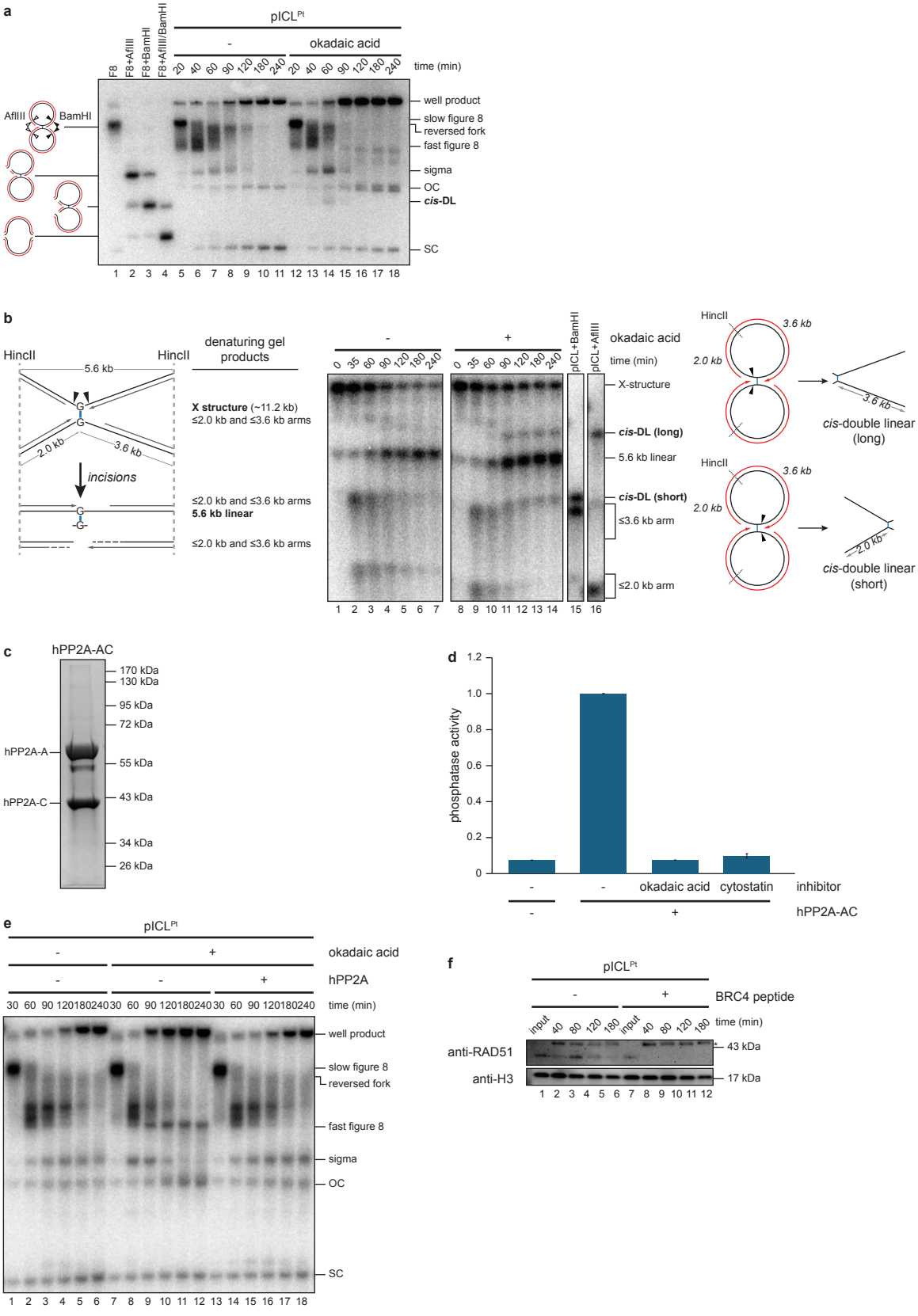
Supplementary Table 3.1 | Oligonucleotide sequences used in this study



Supplementary Fig. 3.1 | Validation of RPE-1 cells used in CRISPR dropout screen

a, Lysates from RPE-1 cells used for the chemogenomic screens described in Fig. 3.1b,c were blotted for FANCD2 and RPA32 (loading control).

b, Relative *FANCD2*$^{+/+}$ and *FANCD2*$^{-/-}$ RPE-1 cell proliferation after psoralen + UVA treatment was measured using CellTiter-Glo reagent. Cell proliferation was measured after 5 days. Data represent the mean ± standard deviation from two independent experiments. FANCD2-deficient RPE-1 cells are hypersensitive to ICLs induced by psoralen + UVA.


a
pICL^Pt
-
okadaic acid
F8
F8+AflIII
F8+BamHI
F8+AflIII/BamHI
20 40 60 90 120 180 240 20 40 60 90 120 180 240
time (min)
AflIII
BamHI
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
cis-DL
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
b
HincII
HincII
5.6 kb
2.0 kb
3.6 kb
incisions
denaturing gel products
X structure (~11.2 kb)
≤2.0 kb and ≤3.6 kb arms
≤2.0 kb and ≤3.6 kb arms
5.6 kb linear
≤2.0 kb and ≤3.6 kb arms
0 35 60 90 120 180 240 0 35 60 90 120 180 240
pICL+BamHI
pICL+AflIII
okadaic acid
time (min)
X-structure
cis-DL (long)
5.6 kb linear
cis-DL (short)
≤3.6 kb arm
≤2.0 kb arm
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
HincII
3.6 kb
2.0 kb
3.6 kb
cis-double linear (long)
HincII
3.6 kb
2.0 kb
2.0 kb
cis-double linear (short)
c
hPP2A-AC
170 kDa
130 kDa
95 kDa
72 kDa
hPP2A-A
55 kDa
hPP2A-C
43 kDa
34 kDa
26 kDa
d
phosphatase activity
1.2
1.0
0.8
0.6
0.4
0.2
0
- - okadaic acid cytostatin inhibitor
- + hPP2A-AC
e
pICL^Pt
- +
okadaic acid
- - +
hPP2A
30 60 90 120 180 240 30 60 90 120 180 240 30 60 90 120 180 240
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
f
pICL^Pt
- +
BRC4 peptide
input 40 80 120 180 input 40 80 120 180
time (min)
anti-RAD51
43 kDa
anti-H3
17 kDa
1 2 3 4 5 6 7 8 9 10 11 12

Supplementary Fig. 3.2 | Characterization of okadaic acid-induced fork collapse products

a, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and okadaic acid, as indicated. Replication intermediates were resolved on a native agarose gel as in Fig. 3.2a. The identities of standards generated by digesting cross-linked figure 8 replication intermediates with AflIII (white arrowheads) and/or BamHI (black arrowheads) are shown to the left of the gel. A single cut by either AflIII or BamHI produces a species that migrates like a sigma intermediate, while two cuts produce a species that migrates similarly to a 10 kilobase pair linear DNA (*cis*-double linear). Digestion with both AflIII and BamHI produces a ~5.6 kilobase pair linear. Note that an okadaic acid induced putative cleavage product comigrates with the *cis*-double linear cleavage standard, suggesting this species is produced upon cleavage of both parental DNA strands at one fork.

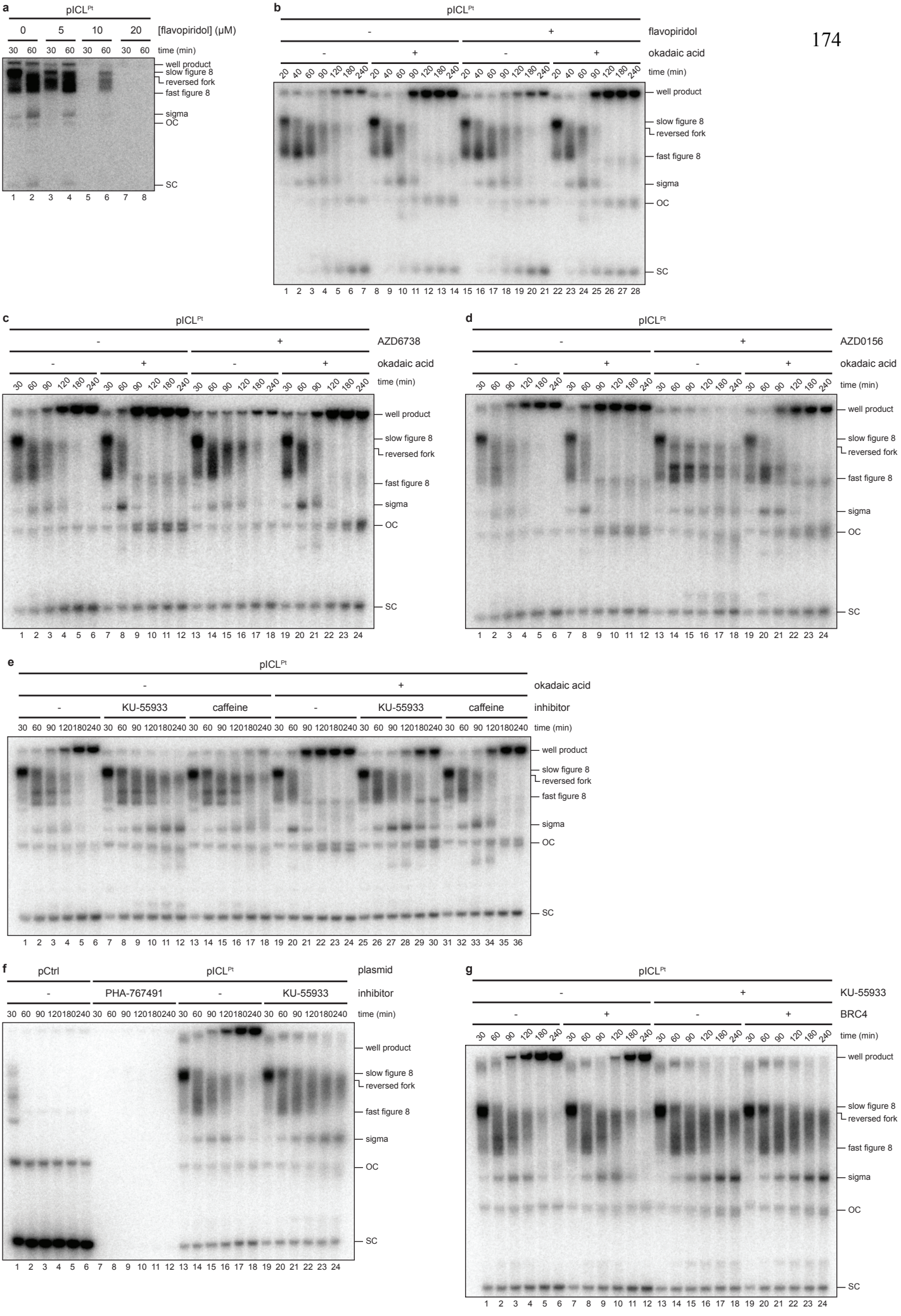
b, Left, schematic of species generated during ICL repair. Digestion with HincII produces X-structures corresponding to cross-linked plasmids and linear species corresponding to unhooked plasmids. Right, pICL^{Pt} was replicated in egg extract supplemented with okadaic acid, as indicated. Recovered replication intermediates were digested with HincII and resolved on a denaturing (alkaline) agarose gel. Intermediates were detected by Southern blotting with ^{32}P radiolabeled probes and visualized by autoradiography. Samples were run alongside unreplicated pICL^{Pt} that was digested with HincII and BamHI or AflIII to generate standards corresponding to cleavage of both parental DNA strands on one side or the other of the ICL (referred to as long or short *cis*-double linear). Note that replication in the presence of okadaic acid results in an accumulation of species that comigrate with both types of *cis*-double linear species, indicating that these species are derived from replication intermediates in which both parental DNA strands in one fork have been cleaved.

c, Purified recombinant hPP2A-AC complex used in **e** was resolved by SDS-PAGE and visualized with Coomassie stain.

d, Phosphatase activity of the hPP2A-AC complex shown in **c** was confirmed using the Serine/Threonine Phosphatase Assay System (Promega). Activity was normalized to the maximum for each experiment. Data represent the mean ± standard deviation from three independent experiments. Note that hPP2A-AC phosphatase activity is suppressed to background levels in the presence of okadaic acid or cytostatin.

e, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and okadaic acid and hPP2A-AC complex, as indicated. Replication intermediates were analyzed as in Fig. 3.2a.

f, pICL^{Pt} was replicated in egg extract supplemented BRC4 peptide, as indicated. Recovered chromatin-associated proteins were blotted for RAD51 and H3 (loading control) as in Fig. 3.2f. Asterisk indicates non-specific band.


a
pICL^Pt
0
5
10
20
[flavopiridol] (μM)
30 60 30 60 30 60 30 60 time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8
b
pICL^Pt
flavopiridol
okadaic acid
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
c
pICL^Pt
AZD6738
okadaic acid
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
d
pICL^Pt
AZD0156
okadaic acid
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
e
pICL^Pt
okadaic acid
KU-55933
caffeine
inhibitor
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36
f
pCtrl
pICL^Pt
plasmid
PHA-767491
KU-55933
inhibitor
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
g
pICL^Pt
KU-55933
BRC4
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

Supplementary Fig. 3.3 | Effects of CDK and DNA damage checkpoint inhibitors on ICL repair by the FA pathway

a, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and pretreated with the pan-CDK inhibitor flavopiridol, as indicated. Replication intermediates were resolved on a native agarose gel as in Fig. 3.2a. Note that treatment with flavopiridol prior to initiation of replication is sufficient to abolish replication, indicating that the concentration of flavopiridol used in subsequent reactions effectively inhibits CDK activity in egg extracts.

b, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and treated with flavopiridol after initiating replication, as indicated. Replication intermediates were resolved on a native agarose gel as in Fig. 3.2a. Flavopiridol treatment has no discernible effect on processing of ICL repair intermediates, either in the presence or absence of okadaic acid.

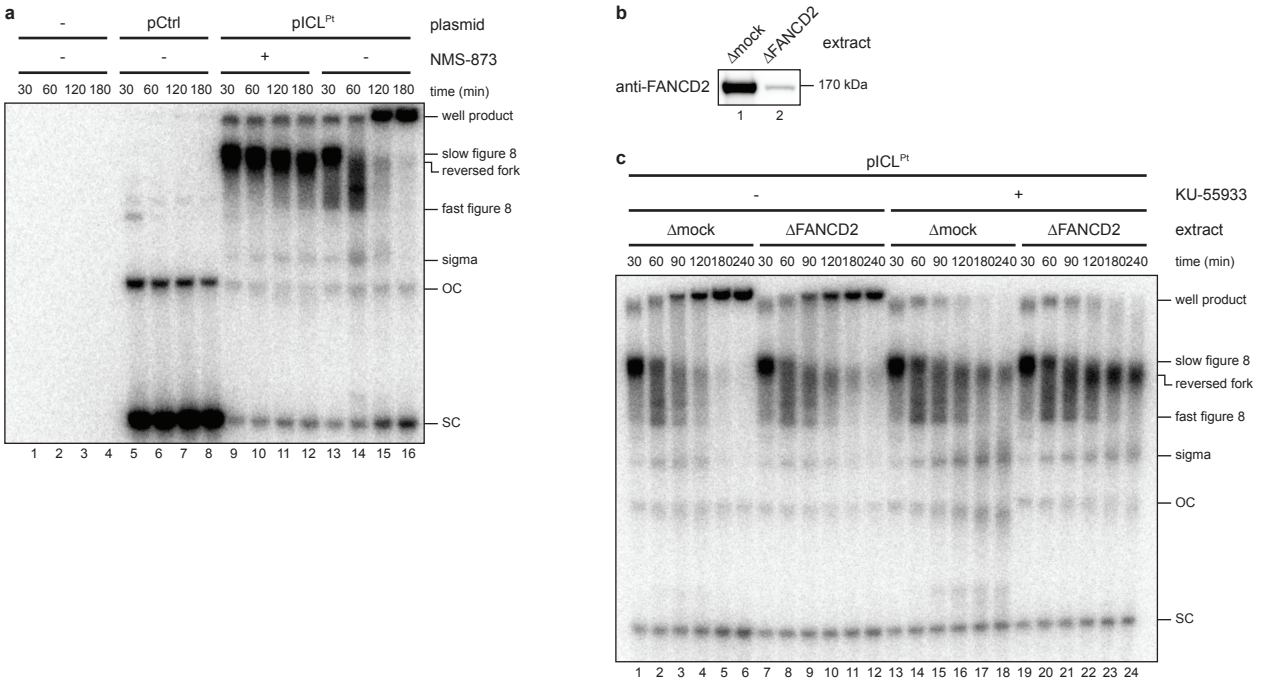
c, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and the ATR inhibitor AZD6738, as indicated. Replication intermediates were resolved on a native agarose gel as in Fig. 3.2a. Treatment with AZD6738 resulted in modest stabilization of reversed fork intermediates and partially suppressed okadaic acid-induced formation of putative fork cleavage products.

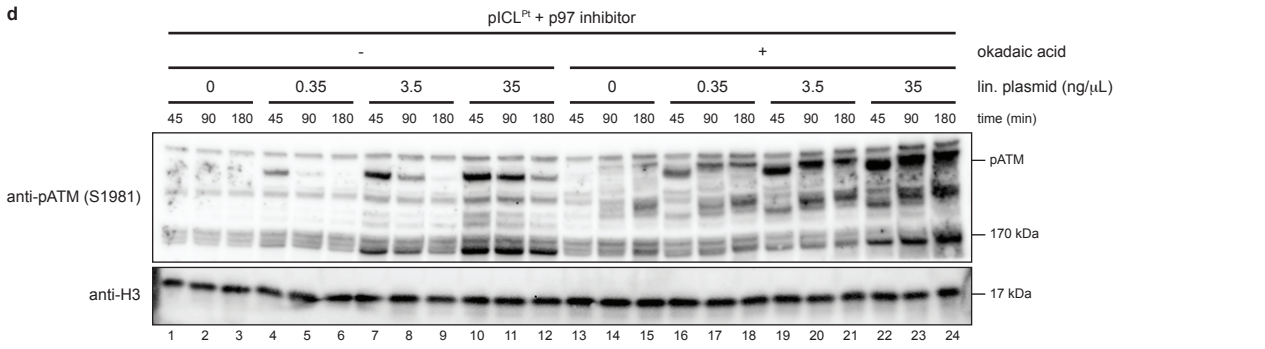
d, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and the ATM inhibitor AZD0156, as indicated. Replication intermediates were resolved on a native agarose gel as in Fig. 3.2a. As with KU-55933 treatment, AZD0156 treatment resulted in stabilization of reversed fork intermediates and suppressed formation of okadaic acid-induced putative fork cleavage products.

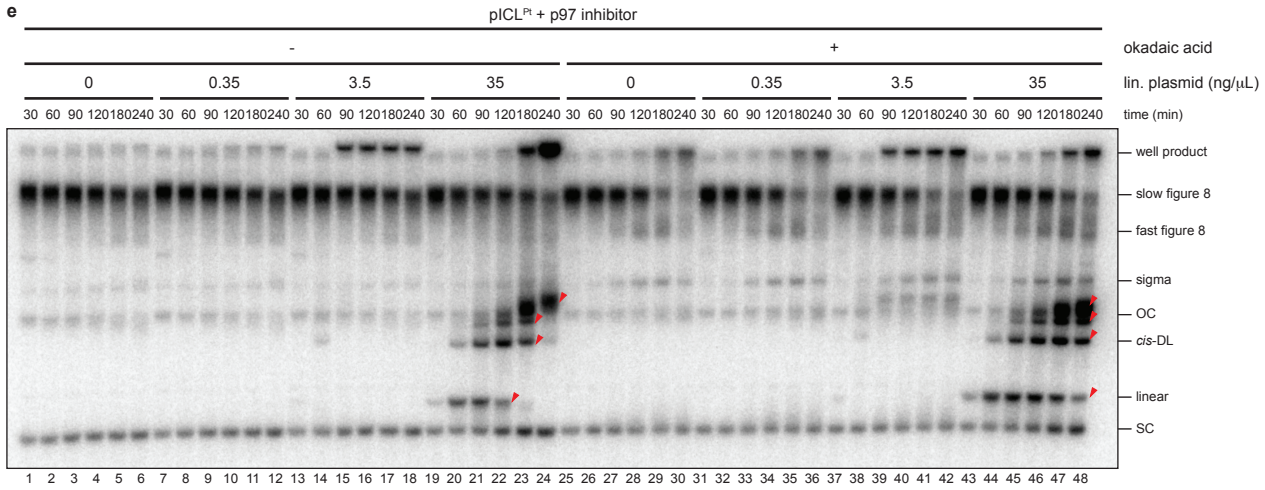
e, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and okadaic acid and the ATM inhibitors KU-55933 or caffeine, as indicated. Replication intermediates were analyzed as in Fig. 3.2a.

f, The replication reactions described in Fig. 3.3f were supplemented with [α-^{32}P]dCTP, and replication intermediates were resolved on a native agarose gel as in Fig. 3.2a. A representative gel corresponding to the samples used for the RPA70 blot is shown.

g, The replication reactions described in Fig. 3.3g were supplemented with [α-^{32}P]dCTP, and replication intermediates were resolved on a native agarose gel as in Fig. 3.2a.


a
plasmid
pCtrl
pICL^Pt
NMS-873
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
b
extract
Δmock
ΔFANCD2
anti-FANCD2
170 kDa
c
pICL^Pt
KU-55933
extract
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC


d
pICL^Pt + p97 inhibitor
okadaic acid
lin. plasmid (ng/μL)
time (min)
anti-pATM (S1981)
pATM
170 kDa
anti-H3
17 kDa


e
pICL^Pt + p97 inhibitor
okadaic acid
lin. plasmid (ng/μL)
time (min)
well product
slow figure 8
fast figure 8
sigma
OC
cis-DL
linear
SC

Supplementary Fig. 3.4 | ATM activation promotes processing of reversed fork intermediates

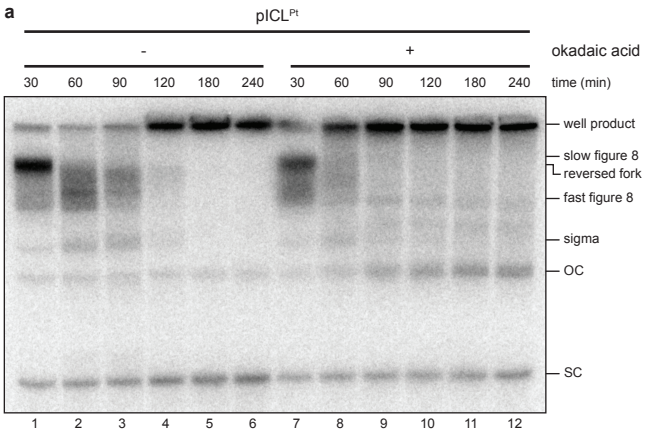
a, The replication reactions described in Fig. 3.4a were supplemented with [α-^{32}P]dCTP, and replication intermediates were resolved on a native agarose gel as in Fig. 3.2a.

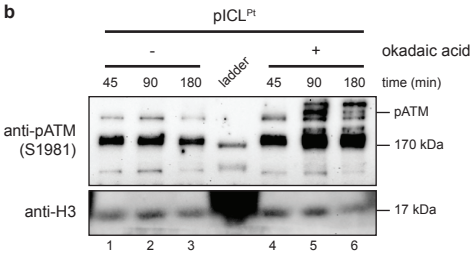
b, FANCD2 immunodepletion. Egg extracts used in Fig. 3.4b were resolved by SDS-PAGE and blotted for FANCD2.

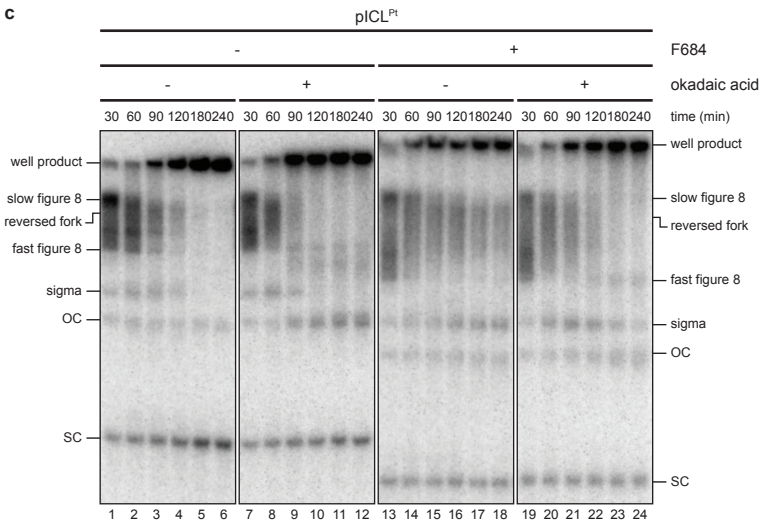
c, The replication reactions described in Fig. 3.4b were supplemented with [α-^{32}P]dCTP, and replication intermediates were resolved on a native agarose gel as in Fig. 3.2a.

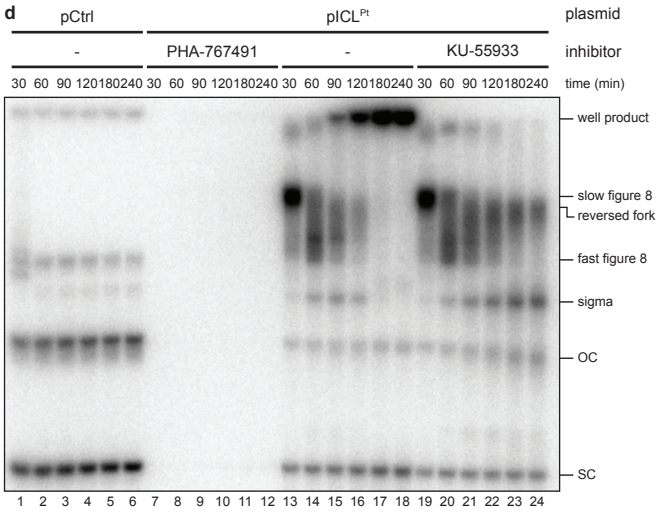
d, pICL^{Pt} was replicated in egg extract supplemented with NMS-873 (to prevent CMG unloading), okadaic acid, and increasing concentrations of linearized pCtrl, as indicated. Replication reactions were blotted for phospho-ATM (S1981) and H3 as in Fig. 3.3b. Addition of linear DNA induces robust ATM activation.

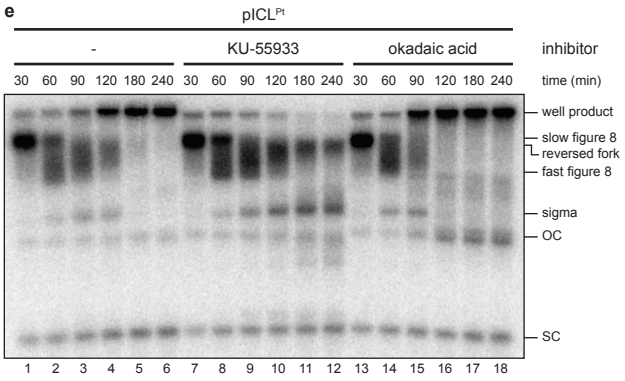
e, The pICL^{Pt} replication reactions described in **d** were additionally supplemented with [α-^{32}P]dCTP, and replication intermediates were resolved on a native agarose gel as in Fig. 3.2a. Despite extensive ATM activation, slow figure 8 intermediates corresponding to converged replication fork structures are largely refractory to resection and fork collapse. Red arrowheads indicate linear plasmids and end-joining products that become labeled due to unscheduled DNA synthesis on the linearized plasmid in extract.


a
pICL^Pt
-
+
okadaic acid
30 60 90 120 180 240 30 60 90 120 180 240
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12


b
pICL^Pt
-
ladder
+
okadaic acid
45 90 180 45 90 180
time (min)
anti-pATM (S1981)
pATM
170 kDa
anti-H3
17 kDa
1 2 3 4 5 6


c
pICL^Pt
-
+
F684
-
+
-
+
okadaic acid
30 60 90 120180240 30 60 90 120180240 30 60 90 120180240 30 60 90 120180240
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24


d
pCtrl
pICL^Pt
plasmid
-
PHA-767491
-
KU-55933
inhibitor
30 60 90 120180240 30 60 90 120180240 30 60 90 120180240 30 60 90 120180240
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24


e
pICL^Pt
-
KU-55933
okadaic acid
inhibitor
30 60 90 120 180 240 30 60 90 120 180 240 30 60 90 120 180 240
time (min)
well product
slow figure 8
reversed fork
fast figure 8
sigma
OC
SC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Supplementary Fig. 3.5 | Nuclease-dependent processing of reversed fork intermediates

a, The untreated and okadaic acid-treated replications analyzed by mass spectrometry in Fig. 3.5a were supplemented with [α-^{32}P]dCTP, and replication intermediates were resolved on a native agarose gel as in Fig. 3.2a.

b, The untreated and okadaic acid-treated replications analyzed by mass spectrometry in Fig. 3.5a were resolved by SDS-PAGE and blotted for phospho-ATM (S1981) and H3 (loading control) to confirm ATM activation.

c, pICL^{Pt} was replicated in egg extract supplemented with [α-^{32}P]dCTP and okadaic acid and the EXO1 inhibitor F684, as indicated. Replication intermediates were analyzed as in Fig. 3.2a. Note that the replication reactions shown were performed in parallel, but lanes 1-12 and lanes 13-24 were resolved on separate native agarose gels, resulting in differences in the migration of replication intermediates and products across gels.

d, The replication reactions described in Fig. 3.5e were supplemented with [α-^{32}P]dCTP, and replication intermediates were resolved on a native agarose gel as in Fig. 3.2a. A representative gel corresponding to the samples used for the DNA2 blot is shown.

e, The replication reactions described in Fig. 3.5f were supplemented with [α-^{32}P]dCTP, and replication intermediates were resolved on a native agarose gel as in Fig. 3.2a.

Chapter 4

REPLICATION-INDEPENDENT ACTIVATION OF THE FANCONI ANEMIA PATHWAY BY DNA REPLICATION INTERMEDIATES

Maria Altshuller[1] and Daniel R. Semlow[1]

Affiliations

[1]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA

4.1 Abstract

DNA interstrand cross-link (ICL) repair by the Fanconi anemia (FA) pathway is coupled to DNA replication: convergence of two replication forks at an ICL triggers a cascade of repair events, including CMG helicase unloading, fork reversal, nucleolytic unhooking, translesion synthesis, and homologous recombination. However, a key unresolved question is whether FA pathway activation requires the replisome itself or whether the DNA structures produced during replication are sufficient to trigger the FA pathway. This distinction that has been difficult to address experimentally, as replication and replication-dependent repair cannot be decoupled in conventional assays. Here, we describe a cell-free experimental system in which X-shaped replication intermediates, which are generated by replicating ICL-containing plasmids, are purified and added to non-replicating *Xenopus* egg extract. We present preliminary data showing that purified X-structures are efficiently processed into a series of intermediates consistent with FA pathway-mediated repair, including reversed forks, post-incision sigma structures, and linear species, ultimately yielding monomeric repair products. We further show that incubation of X-structures in non-replicating extract activates the replication stress response. Together, these findings suggest that X-shaped intermediates may be sufficient to activate ICL repair without the replisome, which points toward a model in which X-shaped intermediates generated by either fork convergence or ICL traverse can serve as a common activator of FA pathway-mediated ICL repair. Furthermore, this work establishes a new biochemical platform for dissecting the structural requirements for FA pathway activation.

4.2 Introduction

A central but unresolved question in the ICL repair field concerns the mechanistic basis for the coupling of FA pathway activation to DNA replication. When a replication fork encounters an ICL, the replisome stalls and initiates the sequence of events characteristic of FA pathway-mediated repair: CMG helicase unloading, fork reversal, ICL unhooking, fork restoration, translesion synthesis, and HR-mediated DSB repair[1]. However, it is unclear whether the replisome itself is required for this activation – for example, it may be needed to target repair factors to the lesion – or whether the DNA structures generated by replication fork convergence are themselves sufficient to engage the repair machinery. This distinction has important implications for elucidating the structural basis of ICL recognition by the FA pathway and potentially for understanding how ICL repair is activated in contexts where the replisome bypasses the lesion rather than stalling at it, as occurs during ICL traverse in mammalian cells.

Addressing this question has been technically challenging because, in conventional replication-coupled ICL repair experiments, the replisome and the DNA structures it generates are inseparable. To overcome this, we developed an approach to uncouple ICL repair from the process of replication by purifying repair intermediates and adding them to non-replicating extract, allowing us to ask whether specific DNA structures are sufficient to activate repair in the absence of replication machinery. Using this system, we obtained proof-of-concept evidence that X-shaped replication intermediates incubated in non-replicating extract can activate ATR signaling and induce FANCD2 monoubiquitylation, which are features consistent with FA pathway activation. Additionally, we found that the

X-structures can be efficiently resolved into repair products through a sequence of intermediates that resembles those observed during FA pathway-mediated repair.

These findings suggest that the DNA X-structures themselves, rather than the replisome, may be the proximal trigger for FA pathway activation. They also raise several important questions that this appendix begins to address: is the processing of these intermediates dependent on canonical FA pathway factors? Which specific DNA structures – X-shaped structures, reversed fork structures, and/or post-incision sigma structures – can serve as substrates for the FA pathway? Which replication factors, if any, need to be recruited to facilitate FA pathway-mediated repair? Although this project is in its early stages, the preliminary data presented here not only establish a new experimental framework that will allow these questions to be explored in future work, but also point toward a model that unifies the canonical fork convergence-dependent FA pathway and ICL traverse: in this model, the common X-shaped structures generated through either process serve as the trigger for ICL resolution by the FA pathway.

4.3 Preliminary Data

We first asked whether a purified DNA structure containing replication forks converged at an ICL would be sufficient to activate the Fanconi anemia pathway. To address this question, we replicated a plasmid containing a site-specific cisplatin-ICL ($pICL^{Pt}$) in *Xenopus* egg extracts in the presence of [α-^{32}P]dCTP (which is incorporated into nascent DNA) and the p97 inhibitor NMS-873[2]. While this type of ICL would normally be repaired by the FA pathway during replication in extract[3,4], addition of p97 inhibitor prevents CMG unloading and leads to the accumulation of structures in which CMG helicases have

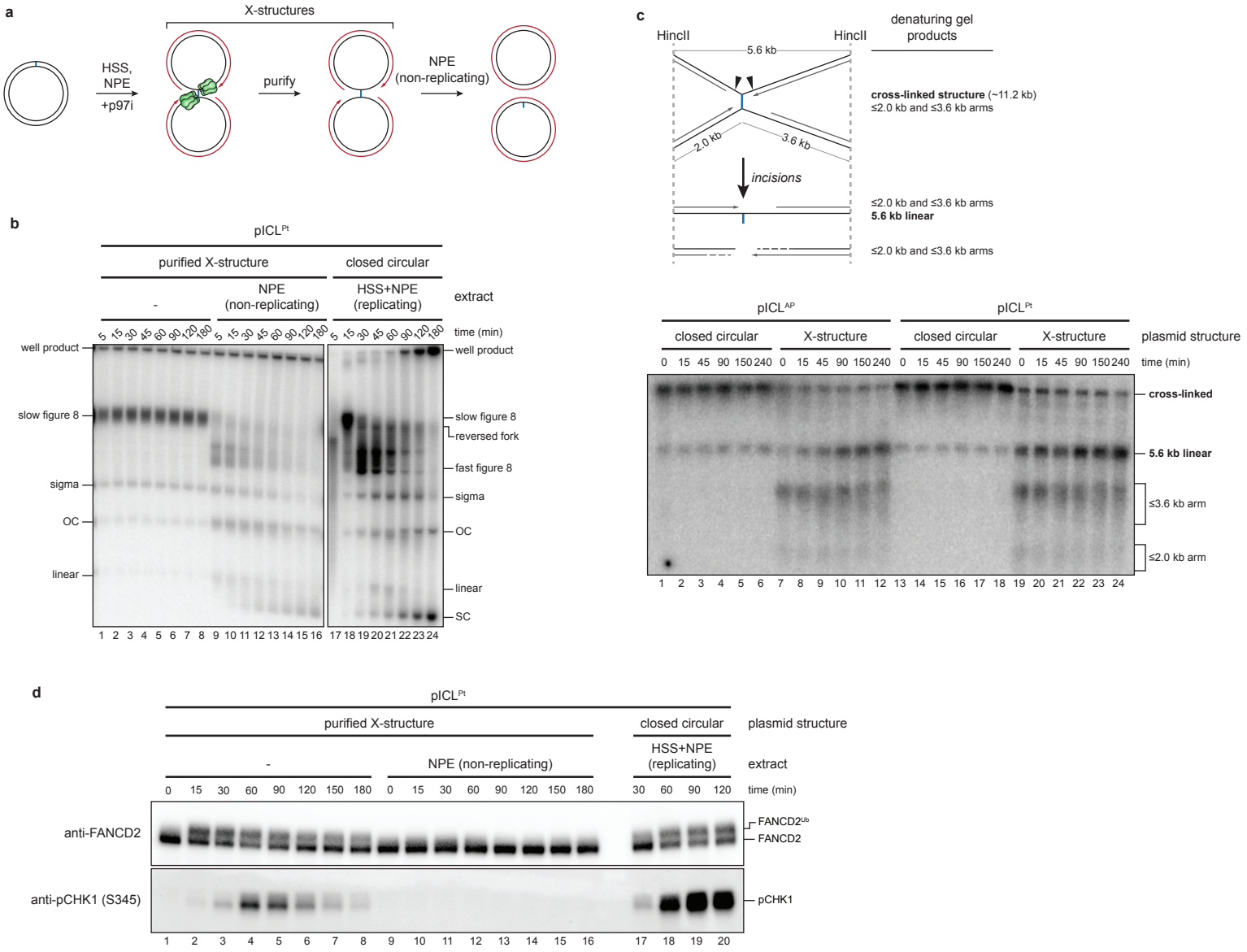


Fig. 4.1 | Purified X-structures undergo ICL resolution and activate FA pathway signaling in non-replicating extract

a, Schematic describing the experimental approach. An ICL-containing plasmid is replicated in *Xenopus* egg extracts supplemented with p97 inhibitor to stall converged replication forks at the ICL. The resulting X-structures are purified and can subsequently be added to non-replicating NPE.

b, pICL^{Pt} X-structures were purified as described in **a** from a replication reaction that also contained [α-^{32}P]dCTP. Purified X-structures were incubated in non-replicating extract alongside a standard pICL^{Pt}

replication, and repair intermediates were recovered, deproteinized, separated on a native agarose gel, and visualized by autoradiography. SC, supercoiled; OC, open circular.

c, Top, schematic of species generated during ICL repair. Digestion with the restriction enzyme HincII produces cross-linked species and linear species that have undergone ICL unhooking. Bottom, X-structures purified from $pICL^{AP}$ and $pICL^{Pt}$ replication reactions were incubated in non-replicating extract, linearized with HincII, and resolved on a denaturing (alkaline) agarose gel. Repair intermediates were detected by Southern blotting and visualized by autoradiography. Note that closed circular plasmids largely remained cross-linked in non-replicating extract, while purified X-structures appeared to be resolved over time.

d, $pICL^{Pt}$ was incubated in extract supplemented with p97 inhibitor to purify X-structures and in extract supplemented with p97 inhibitor and CDC7 inhibitor to purified closed circular species. These purified plasmids were incubated in non-replicating NPE alongside a standard replication reaction of (unpurified) closed circular $pICL^{Pt}$. Samples were resolved by SDS-PAGE, and CHK1 phosphorylation and FANCD2 monoubiquitylation was monitored by immunoblotting.

converged and stalled at the ICL (X-structures). These replication intermediates were purified and added to NPE, which is not sufficient for replication because its high CDK activity prevents loading of MCM2-7 onto chromatin (Fig. 4.1a)[5]. When we resolved the incubation intermediates on a native agarose gel and visualized them by autoradiography, we observed that while the purified X-structures remained stable in the absence of NPE, they were efficiently converted to other species in extract (Fig. 4.1b). Strikingly, the X-structures were processed into a series of intermediates that resembles the pattern of intermediates generated during replication and FA pathway-mediated repair of $pICL^{Pt}$ (Fig. 4.1b, compare lanes 9-16 and 17-24). In particular, the accumulation of fast figure 8 intermediates and reversed forks suggests the occurrence of fork remodeling events, as in the FA pathway, while the accumulation of post-incision sigma structures and linear species is consistent with ICL unhooking. The processing kinetics also appeared consistent with

FA-pathway mediated repair: by the end of the 3-hour time course, most of the intermediates formed during both X-structure incubation and pICL^{Pt} replication had been resolved into open circular and supercoiled repair products (Fig. 4.1b, compare lanes 16 and 24). These data imply that the FA pathway may be responsible for processing X-structures in non-replicating extract, which would mean that activation of the FA pathway can be uncoupled from replication.

To further clarify whether the cross-link in the X-structures can be unhooked in the absence of replication, we replicated plasmids containing two types of ICLs – pICL^{Pt} and pICL^{AP}, an abasic site-ICL that is generally repaired by the NEIL3 pathway – in *Xenopus* egg extracts in the presence of p97i. To minimize the amount of contaminating species, we purified the X-structures from an agarose gel by electroelution and incubated them in non-replicating extract alongside closed circular plasmids. The incubation intermediates were recovered, linearized with a restriction enzyme (HincII), resolved on a denaturing (alkaline) agarose gel, and detected by Southern blotting (Fig. 4.1c). Incubation of the closed circular plasmids led to the accumulation of mostly slow-moving cross-linked structures, which remained largely stable throughout the time course. On the other hand, incubation of both AP-ICL and Pt-ICL X-structures led to the gradual accumulation of unhooked linear species, suggesting that unhooking of the ICL may be occurring in the absence of replication. Notably, if both types of ICLs in this assay are indeed being unhooked, this raises the question of whether this occurs through the same mechanism. Generally, AP-ICLs would be unhooked by the NEIL3 glycosylase rather than the structure-specific nucleases involved in the FA pathway, but whether this holds true for

purified pICL^{AP} X-structures is an interesting open question that this experimental framework will be well-positioned to address in the future.

Since FA pathway activation involves certain hallmarks, such as ATR activation and FANCD2 monoubiquitylation[6-8], we next wanted to determine whether these indicators of possible FA pathway involvement could be observed upon X-structure incubation in the absence of replication. For this assay, we purified X-structures from extract supplemented with p97i (as in Fig. 4.1a) and closed circular species from extract supplemented with p97i and the replication initiation inhibitor CDC7i (PHA-767491)[9], ensuring that the closed circular species underwent the same handling as the X-structures. These purified species were incubated in non-replicating extract, and phosphorylation of CHK1 and monoubiquitylation of FANCD2 were detected by Western blotting (Fig. 4.1d). The closed circular species led to accumulation of neither phospho-CHK1 nor FANCD2-ub; on the other hand, the X-structures induced marked CHK1 phosphorylation and FANCD2 ubiquitylation, both of which are indicative of a replication stress response. Whether the replication stress response pathway being activated by these intermediates is indeed the FA pathway will need to be clarified in future work, but these data suggest the possibility of replication-independent FA pathway activation and establish a biochemical platform for investigating ICL repair in the absence of replication in *Xenopus* egg extracts.

4.4 Future Directions

The preliminary data presented in this appendix represent proof of concept that purification of replication and repair intermediates can be used to study DNA repair uncoupled from replication in *Xenopus* egg extracts. These data further demonstrate that X-shaped

replication intermediates seem to be sufficient to activate features of the FA pathway and that they can be resolved into monomeric repair products in the absence of ongoing DNA replication. These findings motivate a unified model for replication-coupled ICL repair in which X-shaped structures, which can be generated either through replication fork convergence or ICL traverse, serve as the common proximal trigger for FA pathway activation (Fig. 4.2). In fork convergence-dependent repair – such as the canonical FA pathway – the converged forks generate an X-shaped intermediate following CMG unloading; in the traverse pathway, the bypass of the ICL by a single CMG helicase generates an analogous X-shaped structure. According to our proposed model, the downstream repair events – fork reversal, ICL unhooking, fork restoration, TLS, and HR – are then initiated by the X-structure itself rather than by the replisome. Future work will be needed to rigorously test this model and to determine the structural features of the intermediates that engage the FA pathway.

A critical first step will be to determine whether the unhooking of X-shaped intermediates in non-replicating extract depends on canonical FA pathway factors. Although the accumulation of CHK1 phosphorylation and FANCD2 monoubiquitylation upon X-structure incubation is consistent with FA pathway activation, these are not exclusive markers of the FA pathway, and it remains possible that an alternative repair mechanism is responsible for the unhooking activity we observe. To address this, we can immunodeplete FANCI-FANCD2 and the XPF-ERCC1-SLX4 nuclease complex, which is responsible for performing nucleolytic incisions in the FA pathway[10], from extract prior to X-structure incubation. If unhooking is inhibited by the absence of these factors, this would provide further evidence for the FA pathway being able to resolve X-shaped intermediates in the

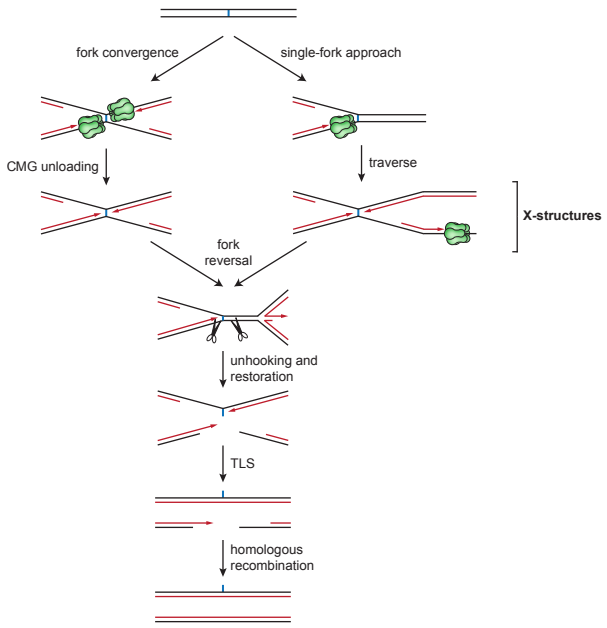


Fig. 4.2 | Model for FA pathway-mediated ICL unhooking following either fork convergence or ICL traverse

Both replication fork convergence followed by CMG unloading (left branch) and single-fork ICL traverse (right branch) generate X-shaped structures. This model proposes that these X-structures can then engage the FA pathway, allowing them to be unhooked through the standard FA pathway mechanism regardless of their origin. See main text for details.

absence of replication. Rescue experiments with recombinant FANCI-FANCD2^{WT} and FANCI-FANCD2^{K562R} (which is FANCD2 monoubiquitylation-dead), both of which have been successfully purified in our laboratory[11], would further elucidate which features of the complex are required for replication-independent repair. Should unhooking prove to be independent of FANCI-FANCD2 and XPF-ERCC1, the involvement of alternative nucleases, such as FAN1, would also be investigated.

In parallel, it will be important to determine which specific DNA structures within the pool of ICL repair intermediates serve as substrates for the FA pathway. Outside of the X-

structures we have worked with up to this point, other structures of interest include Y-shaped structures, which mimic the product of fork reversal and which can be generated by replicating a plasmid containing a LacI-bound *lacO* array positioned to stall one fork while the other encounters the ICL[12,13]. Sigma structures, which represent the product of ICL unhooking, can be enriched for subsequent gel-purification by blocking fork restoration with inhibitors of ATM or DNA2, as shown in Chapter 3. Incubating each of these purified intermediates in non-replicating NPE will allow us to determine whether each structure is independently capable of activating the FA pathway or whether earlier steps in the FA pathway involve recruitment of certain required factors that cannot be engaged otherwise. Furthermore, these purified intermediates can be used to perform plasmid pulldowns coupled to proteomic mass spectrometry[14], which will reveal which FA pathway factors and other repair proteins are recruited to each species. By comparing the repair of X-, Y-, and sigma structures in non-replicating extract, we will gain mechanistic insight into how the ICL is specified for unhooking and which factors are required for each intermediate to transition to the next.

Our preliminary data also raise the interesting question of whether different ICL structures are resolved by different mechanisms in the context of purified intermediates. We observe unhooking of both $pICL^{Pt}$ and $pICL^{AP}$ X-structures in non-replicating extract (Fig 4.1c), but whether both are processed by the FA pathway or by their respective canonical repair mechanisms is unclear. During replication-coupled repair, AP-ICLs are preferentially unhooked by the NEIL3 glycosylase rather than by FA pathway nucleases, though the requirement for fork convergence (and therefore X-structure formation) is a shared feature of both pathways[15,16]. Whether NEIL3 retains its precedence over the FA pathway in the

context of a purified pICL^{AP} X-structure, in which the replisome context that normally accompanies NEIL3-dependent repair is absent, remains an open question. To address this, pICL^{Pt} and pICL^{AP} X-structures can be incubated in FANCD2-depleted extract – if unhooking is abrogated for the cisplatin-ICL but not the AP-ICL, this would suggest that their repair proceeds through their respective canonical pathways. Incubation of pICL^{AP} X-structures in NEIL3-depleted extract can also show whether the pattern of repair intermediates resolved on an agarose gel is dependent on NEIL3. These experiments would begin to reveal whether the hierarchical choice between NEIL3 and FA pathway-mediated unhooking is governed by ICL structure or whether replisome context plays a determining role in pathway choice.

An additional future direction concerns the mutational consequences of replication-independent repair. During replication-coupled ICL repair by the FA pathway, TLS past the ICL remnant introduces characteristic mutational signatures that depend on the chemistry of the specific cross-link – for example, repair of cisplatin-ICLs and colibactin-ICLs each produces distinct mutation spectra, as described in Chapters 2 and 3. If the same mutation signatures are produced during replication-independent repair of purified X-structures, this would provide important evidence that TLS is proceeding through a similar mechanism in both contexts. Next-generation sequencing (NGS) of repair products recovered after X-structure incubation in non-replicating extract would allow a direct comparison of mutation spectra and may further validate the model that FA pathway-mediated repair is faithfully recapitulated in this system. If the mutation spectra differ from the spectra expected after FA pathway-mediated repair of each type of ICL, this approach may allow us to identify

the alternative repair mechanisms involved, as various repair pathways have characteristic mutational signatures that could show up in this NGS experiment.

Finally, the system described here presents a unique opportunity to investigate whether replication factors have roles in FA pathway-mediated repair that are independent of their functions in DNA replication. Because depletion or inhibition of essential replication factors necessarily blocks upstream replication steps in conventional experiments, their potential repair functions have been difficult to study biochemically. The ability to add purified intermediates to non-replicating extract circumvents this limitation. One potential factor of interest is PCNA, which has numerous interaction partners among FA pathway factors: it physically associates with TLS polymerases and the FANCM-FA core complex[17,18], stimulates XPF activity *in vitro*[19], and interacts with candidate unhooking nucleases SNM1A and FAN1[20,21]. PCNA also serves as the recruitment platform for RFWD3, an E3 ubiquitin ligase that promotes turnover of RPA and RAD51 to facilitate homologous recombination during ICL repair[22,23]. Whether unhooking, TLS, and HR during replication-independent cross-link repair depend on PCNA are questions that this system is now well-positioned to address. Additionally, our system may provide the ability to answer an outstanding question in the field concerning the identity of the polymerase responsible for the insertion step of TLS. During FA pathway-mediated ICL repair, TLS proceeds in two steps: first, a polymerase inserts a nucleotide directly opposite the ICL remnant, and then a second polymerase extends the nascent strand past the lesion[3]. While the Pol κ-REV1-Pol ζ complex and Pol η have been implicated in TLS during ICL repair, immunodepletion of these factors has not identified which polymerase performs the insertion step, raising the possibility that a replicative B-family polymerase – Pol δ, Pol ε,

or Pol α – is responsible[24]. Our non-replicating extract system offers a unique opportunity to test this directly: the B-family polymerase inhibitor aphidicolin can be added to the incubation without blocking upstream replication steps, allowing the contribution of replicative polymerases to TLS to be assessed in isolation[25]. If aphidicolin is found to inhibit the insertion step, individual immunodepletion of Pol δ, Pol ε, and Pol α could then identify the specific insertion polymerase. Together, these experiments will exploit the advantages of the non-replicating extract system to resolve long-standing questions about the mechanisms of FA pathway-mediated ICL repair and to establish whether the structural features of replication fork intermediates are sufficient to recapitulate canonical repair in the absence of ongoing DNA replication.

4.5 Methods

4.5.1 Oligonucleotides

Oligonucleotides used for preparation of cross-linked inserts and were ordered from Integrated DNA Technologies, and their sequences can be found in Supplementary Table 2.1, as they are identical to those used in Chapter 2.

4.5.2 Preparation of Xenopus egg extracts

HSS and NPE were prepared exactly as described in Chapters 2 and 3.

4.5.3 Preparation of oligonucleotide duplexes with site-specific interstrand cross-links

The cross-linked oligonucleotides used in this work were prepared exactly as described in Chapter 2.

4.5.4 Preparation of interstrand cross-link containing plasmids (pICL)

$pICL^{AP}$ and $pICL^{Pt}$ were prepared exactly as described in Chapter 2.

4.5.5 Replication reactions

Replication reactions were carried out exactly as described in Chapters 2 and 3. In brief, plasmids were replicated in *Xenopus* egg extracts at a final concentration of 2.5 ng/μL in standard replication reactions and 5 ng/μL in reactions intended for intermediate purification (see below). Where indicated, licensing mix was supplemented with 167-333 nM 3000 Ci/mmol [α-^{32}P]dCTP and NPE mix was supplemented with 50 μM PHA-767491 or 300 μM NMS-873. For native agarose gel analysis, 1 μL of each reaction was quenched into 6 μL replication stop mix (80 mM Tris-HCl [pH 8.0], 8 mM EDTA, 0.13% phosphoric acid, 10% ficoll, 5% sodium dodecyl sulfate [SDS], 0.1% bromophenol blue) at the indicated time points. Samples were worked up, resolved on a 0.8% agarose gel, and visualized by autoradiography as described in Chapter 2.

4.5.6 Purification of X-structures and closed circular species from replication reactions

Replication reactions were carried out at a plasmid concentration of 5 ng/μL. For purification of hot intermediates, licensing mix was supplemented with [α-^{32}P]dCTP as described above. X-structures were purified from reactions supplemented with 300 μM NMS-873, and closed circular species were purified from reaction supplemented with 50 μM PHA-767491 and 300 μM NMS-873. 40 minutes after initiating replication, the entire reactions (generally 40-60 μL volume) were quenched into 10 times their volume of clear stop mix (50 mM Tris-HCl [pH 8.5], 0.5% SDS, 25 mM EDTA). Samples were digested

with 200 µg/mL RNAse A for 30 mins at 37 °C and with 1 mg/mL proteinase K for 60 mins at 37 °C, extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0) and once with chloroform, ethanol precipitated, and resuspended in 10 µL 10 mM Tris-HCl pH 8.5. Experimentally, we found that X-structures purified with this method were relatively pure, with few enough contaminating species to be useful for our purposes. However, where indicated, additional gel purification was performed to obtain purer species. In these assays, ethanol precipitated samples were resuspended in 10 µL replication stop mix instead of 10 mM Tris-HCl pH 8.5, and the replication intermediates were resolved by native agarose gel electrophoresis. DNA was stained with SYBR Gold and visualized with a Blue-Light Transilluminator, and the species of interest were excised and electroeluted in 1X TBE at 7.5 V/cm at 4 °C. Migration of the species of interest out of the gel matrix was monitored with a UV wand. Purified intermediates were then ethanol precipitated from the eluate, resuspended in 5-10 µL Pt-ICL storage buffer (10 mM Tris-HCl [pH 7.4], 10 mM $NaClO_4$) or AP-ICL storage buffer (10 mM Tris-HCl [pH 7.5]), flash-frozen, and stored at -80 °C.

4.5.7 Incubation in non-replicating extract

Purified intermediates were added to 50% NPE mix to a final concentration of 0.2-2.5 ng/µL (depending on the concentration of the purified intermediate stock). For native agarose gel electrophoresis, 1 µL samples were quenched into 6 µL replication stop mix; for Southern blotting, 4 µL samples were quenched into 40 µL clear stop mix; for Western blotting, 1 µL samples were quenched into 9 µL 1X Laemmli loading buffer. Samples were subsequently treated and analyzed as described in the respective sections for these techniques.

4.5.8 Southern blotting

Southern blotting was performed exactly as described in Chapters 2 and 3.

4.5.9 Immunoblotting

Immunoblotting was performed exactly as described in Chapters 2 and 3. In brief, samples in 1X Laemmli loading buffer were resolved on 10% or 4-15% acrylamide gels and transferred to polyvinyl difluoride (PVDF) membranes. Membranes were blocked with 5% nonfat milk in 1X PBS buffer with 0.05% (v/v) Tween 20 (PBST) for 1 hour at room temperature, rinsed 4 times with 1X PBST, and incubated with primary antibodies diluted to an appropriate concentration in 1X PBST at 4 °C overnight with shaking. Rabbit polyclonal anti-phospho-CHK1 (S345; Cell Signaling Technology #2341) was diluted 1:500, and rabbit polyclonal anti-FANCD2 was diluted 1:5,000[3]. The next day, membranes were washed 3 times with 1X PBST and incubated in goat polyclonal anti-rabbit IgG (H+L) conjugated to horseradish peroxidase (Jackson ImmunoResearch #111-035-003) diluted 1:25,000 in 5% nonfat milk in 1X PBST for 30 mins at room temperature. Following the secondary antibody incubation, membranes were washed 3 times with 1X PBST, incubated for 2-4 mins in chemiluminescence substrate (either Thermo Fisher Scientific #34076 or Genesee Scientific #20-300S), and imaged using a ChemiDoc Imaging System (Bio-Rad).

4.6 References

1 Semlow, D. R. & Walter, J. C. Mechanisms of Vertebrate DNA Interstrand Cross-Link Repair. *Annu Rev Biochem* **90**, 107-135 (2021). https://doi.org/10.1146/annurev-biochem-080320-112510

2 Magnaghi, P. *et al.* Covalent and allosteric inhibitors of the ATPase VCP/p97 induce cancer cell death. *Nat Chem Biol* **9**, 548-556 (2013). https://doi.org/10.1038/nchembio.1313

3 Raschle, M. *et al.* Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* **134**, 969-980 (2008). https://doi.org/10.1016/j.cell.2008.08.030

4 Knipscheer, P. *et al.* The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* **326**, 1698-1701 (2009). https://doi.org/10.1126/science.1182372

5 Walter, J., Sun, L. & Newport, J. Regulated chromosomal DNA replication in the absence of a nucleus. *Mol Cell* **1**, 519-529 (1998). https://doi.org/10.1016/s1097-2765(00)80052-0

6 Tan, W., van Twest, S., Murphy, V. J. & Deans, A. J. ATR-Mediated FANCI Phosphorylation Regulates Both Ubiquitination and Deubiquitination of FANCD2. *Front Cell Dev Biol* **8**, 2 (2020). https://doi.org/10.3389/fcell.2020.00002

7 Kupculak, M. *et al.* Phosphorylation by ATR triggers FANCD2 chromatin loading and activates the Fanconi anemia pathway. *Cell Rep* **42**, 112721 (2023). https://doi.org/10.1016/j.celrep.2023.112721

8 Alcón, P. *et al.* FANCD2-FANCI is a clamp stabilized on DNA by monoubiquitination of FANCD2 during DNA repair. *Nat Struct Mol Biol* **27**, 240-248 (2020). https://doi.org/10.1038/s41594-020-0380-1

9 Montagnoli, A. *et al.* A Cdc7 kinase inhibitor restricts initiation of DNA replication and has antitumor activity. *Nat Chem Biol* **4**, 357-365 (2008). https://doi.org/10.1038/nchembio.90

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). https://doi.org/10.1016/j.molcel.2014.03.015

11 Altshuller, M. *et al.* The Fanconi anemia pathway repairs colibactin-induced DNA interstrand cross-links. *Nat Commun* **16**, 10563 (2025). https://doi.org/10.1038/s41467-025-65606-1

12 Dewar, J. M., Budzowska, M. & Walter, J. C. The mechanism of DNA replication termination in vertebrates. *Nature* **525**, 345-350 (2015). https://doi.org/10.1038/nature14887

13 Zhang, J. *et al.* DNA interstrand cross-link repair requires replication-fork convergence. *Nat Struct Mol Biol* **22**, 242-247 (2015). https://doi.org/10.1038/nsmb.2956

14 Räschle, M. *et al.* DNA repair. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. *Science* **348**, 1253671 (2015). https://doi.org/10.1126/science.1253671

15 Semlow, D. R., Zhang, J., Budzowska, M., Drohat, A. C. & Walter, J. C. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase. *Cell* **167**, 498-511 e414 (2016). https://doi.org/10.1016/j.cell.2016.09.008

16 Wu, R. A. *et al.* TRAIP is a master regulator of DNA interstrand crosslink repair. *Nature* **567**, 267-272 (2019). https://doi.org/10.1038/s41586-019-1002-0

17 Rohleder, F. *et al.* FANCM interacts with PCNA to promote replication traverse of DNA interstrand crosslinks. *Nucleic Acids Res* **44**, 3219-3232 (2016). https://doi.org/10.1093/nar/gkw037

18 Geng, L., Huntoon, C. J. & Karnitz, L. M. RAD18-mediated ubiquitination of PCNA activates the Fanconi anemia DNA repair network. *J Cell Biol* **191**, 249-257 (2010). https://doi.org/10.1083/jcb.201005101

19 Hutton, R. D., Craggs, T. D., White, M. F. & Penedo, J. C. PCNA and XPF cooperate to distort DNA substrates. *Nucleic Acids Res* **38**, 1664-1675 (2010). https://doi.org/10.1093/nar/gkp1104

20 Yang, K., Moldovan, G. L. & D'Andrea, A. D. RAD18-dependent recruitment of SNM1A to DNA repair complexes by a ubiquitin-binding zinc finger. *J Biol Chem* **285**, 19085-19091 (2010). https://doi.org/10.1074/jbc.M109.100032

21 Porro, A. *et al.* FAN1 interaction with ubiquitylated PCNA alleviates replication stress and preserves genomic integrity independently of BRCA2. *Nat Commun* **8**, 1073 (2017). https://doi.org/10.1038/s41467-017-01074-6

22 Feeney, L. *et al.* RPA-Mediated Recruitment of the E3 Ligase RFWD3 Is Vital for Interstrand Crosslink Repair and Human Health. *Mol Cell* **66**, 610-621 e614 (2017). https://doi.org/10.1016/j.molcel.2017.04.021

23 Lin, Y. C. *et al.* PCNA-mediated stabilization of E3 ligase RFWD3 at the replication fork is essential for DNA replication. *Proc Natl Acad Sci U S A* **115**, 13282-13287 (2018). https://doi.org/10.1073/pnas.1814521115

24 Budzowska, M., Graham, T. G., Sobeck, A., Waga, S. & Walter, J. C. Regulation of the Rev1-pol ζ complex during bypass of a DNA interstrand cross-link. *EMBO J* **34**, 1971-1985 (2015). https://doi.org/10.15252/embj.201490878

25 Baranovskiy, A. G. *et al.* Structural basis for inhibition of DNA replication by aphidicolin. *Nucleic Acids Res* **42**, 14013-14021 (2014). https://doi.org/10.1093/nar/gku1209