Roadblocks & bypasses : protection of genome stability by translesion DNA synthesis in C. elegans

Bostelen, I. van

5

A novel assay to study mutation induction of interstrand crosslink repair in *C. elegans*

Ivo van Bostelen, Robin van Schendel, Dennis Brugman & Marcel Tijsterman

Abstract

Interstrand crosslinks (ICLs) are covalent bonds between bases of complementary DNA strands. ICLs are absolute blocks for DNA transcription and replication as they prevent strand separation. These lesions are not only proven cytotoxic, they also have the potential to induce mutations, which range from simple base substitutions to complex genomic rearrangements. One of the most successful systems to unravel a potentially universal mechanism of crosslink repair is the Xenopus egg extract, in which replication and repair of plasmids that contain a single DNA crosslink can be monitored at the molecular level. Ideally, this *in vitro* system is complemented by analogous *in vivo* experiments. We set out to develop such an assay, which is here presented. We demonstrate that the *C. elegans* germline can be used to monitor the repair of the ICL-containing plasmids (pICL) that were previously used in Xenopus egg extracts. We use this novel method to assay different DNA repair deficient backgrounds. We find error free repair and bypass to be affected by defects in nucleotide excision repair (XPA) and translesion synthesis (Polζ, Polη, and PCNA), while mutagenic outcomes are, in part, dependent on polymerase theta-mediated end joining. The established role of these factors in the DNA damage response argues for the validity of this new assay that provides new opportunities to study ICL repair mechanisms at the nucleotide level *in vivo*.

Introduction

Interstrand crosslinks (ICLs) are one of the most toxic types of DNA damage and are potentially very mutagenic. The crosslinking of paired DNA strands through a covalent bond between bases in opposing strands forms an absolute block for transcription and replication because it prevents strand separation. The strong toxicity of ICLs is illustrated by the effective treatment of cancer with agents that induce ICLs - especially replicating cells have great difficulty with ICLs - but also explains why nitrogen mustard has been a very powerful chemical weapon [1]. ICLs are induced by a variety of man-made chemicals such as cisplatin and nitrogen mustards. Although naturally occurring sources of ICLs are very rare they may arise from i) by-products of lipid peroxidation within the cell, specifically aldehydes, ii) the presence of abasic sites, or iii) exposure to natural psoralens [2]. When nitrogen mustard was used as the first chemotherapeutic in 1946 to treat cancer its mode of action was still completely unknown and it took decades before scientists started to understand how cells respond to ICLs and how repair of these lesions occurs [3]. Studies of the genetic disorder Fanconi anemia (FA) have provided great insight in the molecular factors that are involved in crosslink repair. Patients with this syndrome have - among other symptoms - progressive bone marrow failure and greatly increased cancer risk [4]. Moreover, FA cells are hypersensitive to ICL-inducing drugs [5]. Today, nineteen FANC genes are identified in patients. All gene products act in the same ICL repair pathway. Seven of them, FANCA, B, C, E, F, G, and L, constitute the FA core complex that has ubiquitin ligase activity by means of the FANCL subunit [6]. The core complex interacts with FANCM, a structure specific DNA binding protein, and when ICLs are recognized during S-phase the core complex ubiquitinates a complex of two other FANC proteins: FANCD2-FANCI. Ubiquitylated FANCD2/I is stabilized on DNA and promotes both checkpoint activation and the recruitment of repair factors that mediate ICL repair [7,8].

Detailed knowledge about the molecular steps that resolve crosslinks during DNA replication has come from studying the replication of plasmids that contain a single cisplatin crosslink in *Xenopus laevis* egg extracts. Studies from the Walter lab have revealed a replication-dependent mechanisms of ICL repair where two replication forks converge on a single cisplatin crosslink [9]. For this pathway it was shown that FANCD2-FANCI ubiquitination controls the incision steps necessary for ICL unhooking [10]. This generates two substrates for different repair reactions: on the one hand a gapped structure that contains a replication blocking lesion, on the other hand a DNA double strand break. While the gapped substrate requires TLS (to bypass the lesion) and NER (to remove the lesion), the DSB is repaired later via Rad51-dependent recombination [11]. Although it is not completely understood which polymerases play a role in the insertion of nucleotides directly opposite a lesion, it has become clear that in Xenopus egg extracts the TLS polymerases REV1 and Polζ collaborate to extend the nascent DNA strand after

this insertion step [9]. In addition, sequencing of repair products of replication-coupled ICL repair has shown a REV1 dependency for mutagenic products [12].

While it was first thought that this mechanism is the predominant pathway for any ICL, a recent study showed that for psoralen crosslinks and crosslinks originating from abasic sites another mechanism is more relevant, one that avoids the generation of a DSB. Here, the DNA backbone is not incised but one of the two N-glycosyl bonds that form the crosslink is cleaved by the NEIL3 DNA glycosylase. The two resulting daughter molecules have single stranded DNA gaps that are subsequently filled by REV1 mediated TLS. This pathway, which is independent of FAND2-FANCI, highlights that ICL repair is flexible and dependent on the structure of the crosslink [13]. FA factors are believed to be specific for ICL repair that occurs when DNA is being replicated and will not act outside this context. The recognition and repair of ICLs independent of replication is not fully understood but studies have pointed towards nucleotide excision repair (NER) factors and DNA mismatch repair machinery in the recognition of ICLs and in the initiation of repair. The efficiency with which ICLs are repaired outside S-phase is dependent on the level in which they distort the DNA helix structure; lesions that strongly disrupt the normal structure of the DNA helix, like cisplatin crosslinks, are more readily recognized and therefore also removed more efficiently [14]. A specific role might be reserved for XPA in the recognition or stabilization of repair intermediates, especially in cisplatin crosslinks [15-17]. TLS is important in both replication dependent and independent ICL repair but the mechanisms may differ [18,19]. Roles of REV1, Polζ and Polη are not entirely clear, but Polκ and PCNA-K164 ubiquitination functions in replication-independent repair in Xenopus egg extracts and a similar role for Polκ was confirmed in mammalian cells [20].

The cell-free model system of Xenopus has provided many valuable insights in the mechanism of ICL repair, but the generality of it is yet unclear, so are many questions yet unanswered and potentially difficult to answer *in vitro*. Here, we employ the model organism *C. elegans* and a newly designed assay to study how ICLs are repaired and mutations are induced during cell divisions and embryogenesis. We are especially interested in the dual roles of TLS in both the induction of and protection against mutation induction. The *C. elegans* model has previously allowed us to study the role of TLS in response to exogenous and endogenous DNA damage in the form of mono-adducts and UV-induced damage [21-23]. The strong conservation of DNA repair activities also make *C. elegans* well suited to study ICL repair in the context of the DNA damage response [24,25]. The mutagenicity of ICL-inducing agents has been studied in different genetic backgrounds in *C. elegans*. Exposure to mechlorethamine, which induces mostly non-bulky guanine mono-adducts but also 5-7% ICLs [26], induces a modest increase in base substitutions but a marked increase in the frequency of insertions, deletions and other structural rearrangements [27]. Exposure to cisplatin,

which mainly induces intrastrand crosslinks and a small percentage of ICLs (<5%) [28], causes a substantial increase in base substitutions, which is more pronounced in *xpf-1* mutants, as expected from a role for the encoded protein in both NER and ICL repair. Cisplatin also induce deletions (especially in the size range of 3-20 bp), insertions and other rearrangements. The deletion junction characteristics argue for the involvement of an alternative end-joining mechanism that uses microhomology to promote repair. In addition, complex rearrangements were found after exposure to cisplatin and mechlorethamine consistent with persistent replication fork stalling and subsequent DSB formation [27]. Photo-activated psoralens (UV/TMP) also generate ICLs and this agent has been widely used to generate knockouts in *C. elegans* [29,30]. In depth analysis of large numbers of UV/TMP-generated deletion alleles in *C. elegans* has shown that the DSB generated after exposure are repaired via Polθ dependent alternative end-joining [31]. Finally, exposure to the ICL-inducing agent mitomycin C (MMC) did not affect base substitutions rates but instead caused deletions. Approximately 50% MMC deletions are small (<20 bp), but the size range extends to >300Mb [32]. Most, if not all of these chemicals, that produce categorically different types of mutations do not only induce ICLs but also mono-adducts or intra-strand crosslinks that are also substrates for TLS. In order to discriminate the type of lesions that is responsible for the different mutational outcomes we developed an *in vivo* assay that monitors repair of one well-defined ICL in *C. elegans* germ cells.

In this assay, we made use of the fact that transgenes can be efficiently introduced in *C. elegans* by microinjection into the gonadal syncytium of a young adult hermaphrodite (P0 generation). Then, during early embryogenesis, multiple copies of the exogenous DNA concatenate to form a high molecular weight extrachromosomal array. A subset of the next generation (F1) will carry this array and a subset of these transgenic F1 animals will produce transgenic offspring, thus providing a stable transgenic line (TGL) [33]. We used this biology to monitor the fate of injected plasmids that contain one ICL. Transgenic animals that contain the DNA surrounding the ICL were analyzed to determine the repair footprints at the nucleotide level. Here we present the methodology and provide preliminary results obtained in different DNA damage response deficient strains.

Materials & Methods

General culturing and strains used

All strains were cultured according to standard methods as described in [34]. The N2 Bristol strain was used as WT control. The strains with alleles *rev-1(gk455794)*, *rev-3(gk919715)*, *polh-1(ok3317)*, *xpa-1(ok0698)*, *fcd-2(tm1298)* were obtained from the Caenorhabditis Genetics Center, Minnesota, USA. The *polh-1(lf0031)*, *polk-1(lf0029)* and *pcn-1(K165R)* alleles were engineered in our laboratory [21 & chapter 4 of this thesis].

Cisplatin and UV/TMP survival assays

To measure germline sensitivity to photo-activated psoralen, staged L4 animals were first treated with TMP (Sigma-Aldrich, T6137, stock: 2,5 mg/ml in acetone) at a concentration of 10 µg/ml in M9 buffer for 1 hour at RT on a rotor. Animals were then transferred to unseeded 6 cm NGM plates; ~200 animals per dose of UV-A (source: predominantly 366 nm, GE lighting F8T5 BLB U.S.A.). The irradiance of our source was determined using a Blak-ray® long wave ultraviolet meter (model: J221, ser. #12994). Measurements varied slightly between experiments with an average of 140 (±10) µW*cm-2 (equals 1,40 J*m-2*s-1). The exact dose was determined by varying the exposure time. Per dose and genotype 12 exposed animals were transferred and equally divided over four fresh OP50 seeded 6 cm NGM plates and allowed to produce offspring for 48 hours. Subsequently adults were discarded and the brood on the plate was allowed to hatch. 24 hours later the numbers of non-hatched eggs and surviving progeny were determined.

Cisplatin germline sensitivity was performed as follows. Staged L4 populations were soaked for 3 hours in M9 containing cisplatin (Accord Healthcare BV, 1 mg/ml) at indicated doses. After treatment for each dose 4 plates (6 cm, NGM seeded with OP50) with 3 L4 stage animals were allowed to produce offspring for 48 hours at 20°C. Subsequently adults were discarded and the brood on the plate was allowed to hatch. 24 hours later non-hatched eggs and surviving progeny were quantified.

Micro-injections of control, pICL and phenotypic markers

Plasmid injections were performed according to standard *C. elegans* microinjection procedures. In short: L4 animals were picked to fresh OP50 seeded NGM plates and incubated at 15°C, 20-24h pre-injection. Plasmid solutions were injected into the gonadal syncytium of the young adult animals (generation P0). pICL and the control plasmid (pCON, same sequence but without crosslink) are the plasmids described in [9] and were supplied to us by Dr. Puck Knipscheer. Two injection mixes were made. Control injection mix: 10 ng/µl pCON with phenotypic markers 100 ng/µl pRF4 (rol-6(su1006)), 10 ng/µl pGH8 (Prab-3::mCherry::unc-54utr), 2.5 ng/µl 1 pCFJ90 (Pmyo-2::mCherry::unc-54utr) and 5 ng/µl pCFJ104 (Pmyo-3::mCherry::unc-54) [mCherry plasmids are described in 35] in sterile ultrapure ddH2O, and ICL injection mix: 10 ng/µl pICL with phenotypic markers 100 ng/µl pRF4 (rol-6(su1006)), 10 ng/µl pGH8 (Prab-3::mCherry::unc-54utr), 2.5 ng/µl 1 pCFJ90 (Pmyo-2::mCherry::unc-54utr) and 5 ng/µl pCFJ104 (Pmyo-3::mCherry::unc-54) in sterile ultrapure ddH2O.

Selection of transgenic F1 and transgenic lines & making worm lysates

Progeny animals (F1 generation) that expressed mCherry were singled to new plates 3–4 days post injection and allowed reproduce for 48 hours. After generating progeny, single F1's were lysed in 15 µl SWLB (50mM KCL, 10mM Tris-HCL pH 8.3, 2.5 mM MgCl2.6H2O,

0,45% NP40, 0,45% Tween20) and heated for 60°C for 1 hour and 90°C for 15 minutes. The progeny of the F1's, the F2 generation, was screened for mCherry expression too. When F2's expressed mCherry, this was identified a transgenic line (TGL) and 5 mCherry positive F2's are lysed together in one reaction in the same way as the F1 animals.

PCR reaction on pICL, SapI digest & gel electrophoresis

In order to amplify the sequences surrounding the site of the crosslink we performed a nested PCR to achieve high specificity and yield using GoTaq© G2 DNA Polymerase (Promega). For the first (external) PCR we used forward primer ATGCCCTGGCTCACAAATAC, and reverse primer AACCGTATTACCGCCTTTGA, which will produce a product of 1048 bp in the case of repair of the ICL without insertions or deletions. For this external PCR 1.0 µl lysis (from F1 or TGL) was used in a reaction mix of 13.8 µl sqH2O, 0.8 µl Forward primer, 0.8 µl Rev primer, 2.0 µl GoTaq White 5x PCR Buffer (Promega), 2.0 µl GoTaq Green 5x PCR buffer (Promega), 0.4 µl dNTPs 10 mM and 0.2 µl GoTaq-Polymerase 5u/µl (Promega) and PCR program 3 min at 95°C; 35 x (20 sec at 95°C, 30 sec at 54°C, 1 min at 72°C); 3 min at 72°C. For internal PCR 1.0 µl PCR product from external PCR was used as substrate, with the forward primer GACATATGGGAGGGCAAATC and reverse primer AATTTGTGATGCTATTGCTTTATTTG to generate a product of 889bp. Internal PCR reaction: 13.8 µl sqH2O, 0.8 µl forward primer, 0.8 µl reverse primer, 2.0 µl GoTaq White 5x PCR Buffer (Promega), 2.0 µl GoTaq Green 5x PCR buffer (Promega), 0.4 µl dNTPs 10 mM and 0.2 µl GoTaq-Polymerase 5u/µl (Promega). The following PCR program was used for the internal PCR: 3 min. 95°C; 35 x (20 sec at 95°C, 30 sec at 54°C, 1 min at 72°C); 3 min at 72°C. Final PCR products were then analyzed by gel electrophoresis to assess presence of PCR products and the presence of larger deletions/insertions. When the crosslink was repaired error free a restriction site will be present for SapI. To analyze the level of mutation induction internal PCR products were digested by SapI (#R0569L, NE Biolabs) and analyzed by gel electrophoresis to estimate error-free repaired and resistant (mutated) fractions.

Pooling and cloning pICL PCR products

Since each F1 or TGL likely has more than one copy of pICL (or pCON) in their extrachromosomal array the PCR products needed to be cloned to be able to analyze single repair products on nucleotide level with Sanger sequencing. PCR products of approximately 20 F1's were pooled and ligated into pGEM-T following manufacturer's instructions. For the PCR products of the TGLs, 4 or 5 reactions were pooled and ligated into pGEM-T according to manufacturer's protocol. Ligation mix of each pool was used in one transformation reaction in competent *E. coli* DH5α and subsequently 90% and 10% of the reaction was cultured on two selective AXI plates (Ampicillin 100 µg/ml, X-gal 5.0g/ml and IPTG 0,5 mM) for blue/white screening.

Colony PCR

White colonies were picked into sterile ddH2O and incubated at RT for 1h on a shaker. The PCR reaction was performed using 1.0 µl from the colony-ddH2O mix. Forward primer: GTAAAACGACGGCCAG and reverse primer: CAGGAAACAGCTATGAC were used in a mix of 13.8 µl sqH2O, 2.0 µl GoTaq White 5x PCR Buffer (Promega), 2.0 µl GoTaq Green 5x PCR buffer (Promega), 0.4 µl dNTPs 10 mM and 0.2 µl GoTaq-Polymerase 5u/µl (Promega). The following PCR program was used: 3 min at 95°C; 35 x (20 sec at 95°C, 30 sec at 48°C, 1 min at 72°C); 3 min at 72°C. This generates a product of 772bp in the case of repair of the ICL without insertions or deletions. PCR products were analyzed by gel electrophoreses, to confirm proper product yield before Sanger sequencing.

Sanger sequencing and analysis of pICL repair products

For Sanger sequencing of the ICL repair products the sequence primer GTAAAACGACGGCCAG was used. Sequences were mapped to the original sequence of the pICL plasmid with the Sequence Analyzer 8 program developed in our laboratory. For detailed analysis and mapping of complex deletions with insertions to the original pICL sequence the online BLAST tool was used [36].

Results & Discussion

TLS and NER deficiency cause sensitivity towards crosslinking-inducing agents

Although potentially mutagenic, TLS of mono-adducts protects the genome from more severe genomic insults such as deletions that arise when DNA replication is permanently blocked and DNA double strand breaks arise as a consequence. Previously, we have reported on the requirement for functional TLS to maintain a stable genome in *C. elegans* and on the requirement for Polθ-dependent end-joining in case TLS is dysfunctional [21-23]. Data obtained in other systems demonstrated a role for the TLS polymerases REV1 and Polζ in replication-associated ICL repair [13,18]. Polη has been implicated because Polη deficient animals are hypersensitive to cisplatin, an ICL-inducing agent [21], and Polκ, as well as the NER factor XPA, has been suggested to act in a replication-independent mode of repair. [15-17,20].

To provide a context for studying ICL repair in *C. elegans* in molecular detail, we first tested the involvement of TLS factors by exposing young adults of different TLS deficient backgrounds to ICL-inducing agents and quantified the survival of their offspring. Following cisplatin exposure, the most pronounced sensitivity is observed for TLS mutants *polh-1* and *rev-1*: at a dose of 200 µM, which is not toxic for WT animals, we observe almost complete embryonic lethality (Fig.1). At this dose we do not observe any significant sensitivity for the other two TLS polymerase mutants *polk-1* and *rev-3* (REV-3 is the catalytic subunit of Polζ). Especially, the absence of sensitivity of *rev-3*

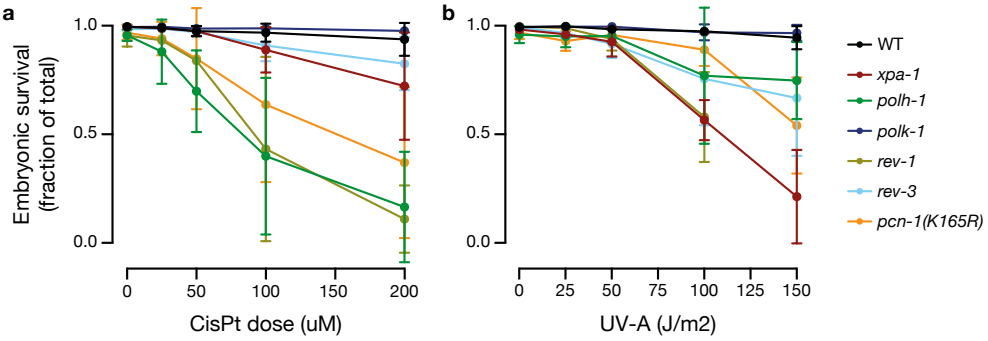


Figure 1. TLS mutants are hypersensitive to crosslinking agents. *Hermaphrodite adult animals of indicated genotypes were exposed to different doses of Cisplatin (a) or UV-TMP (b) and the embryonic survival of the progeny, as a fraction of the total brood, was determined for a 20h time period post exposure.*

mutant animals was surprising: Budzowska *et al.* have described that REV1 performs extension across cisplatin lesions yet in close collaboration with Polζ (thus REV3) [12]. While species-specific mechanistic changes can be argued, an alternative interpretation is that *C. elegans* sensitivity to crosslink-inducing agents do not necessarily read out ICL repair, but instead, translesion synthesis activity across mono-adducts or intrastrand crosslinks - cisplatin induces many more mono-adducts than ICLs: <5% of lesions are ICLs [37]. The sensitivity of *polh-1* mutants may also be seen in that context: while Polη is the most versatile of TLS polymerases and could be a good candidate for the insertion step directly opposite the unhooked crosslink in ICL repair, it likely also is the TLS polymerase to bypass mono-adducts. In keeping with this notion, we find that modification of PCNA at K165 makes worms more sensitive to cisplatin, which augments elaborate data obtained in other species that mono-ubiquitination of PCNA at K165 is a central step in activation of TLS. Yet, *pcn-1(K164R)* animals are not as sensitive to cisplatin as *polh-1* or *rev-1* mutant animals arguing that mono-ubiquination is not essential, or at least not for all lesions. We found no immediate genetic indications for profound replication-independent repair in *C. elegans* as loss of Polκ and XPA does not (or very moderately) sensitize animals to cisplatin exposure in this assay.

We next tested sensitivity of animals exposed to UV-A/TMP. UV-A/TMP induces relatively more crosslinks (*i.e.* psoralens) than cisplatin: up to 40% of all lesions [38]. Here, we did not observe any sensitivity in *polk-1* mutants, intermediate sensitivity in the *polh-1, rev-3* and *pcn-1(K165R)* single mutants and the strongest sensitivity in *xpa-1* and *rev-1* mutants. These results, that are very different from the cisplatin data make clear that animal sensitivity assays without further context is very limited with respect to providing insight into the mechanism of ICL repair. All known ICL-inducing agents also induce mono-adducts and many proteins involved in ICL repair are often

also involved in the response to those. Animal or cell sensitivity can be very useful to establish complementation groups, to categorize pathways, but to address the question on to which lesions these pathways act more specific assays are required.

A novel assay to establish mutagenicity of a single cisplatin crosslink in *C. elegans*

One of the best options to study replication and repair of ICLs is to assay synthetic plasmids that carry a single cisplatin crosslink. This has been done in cell free Xenopus egg extracts, a widely-used model system to study the dynamics of DNA replication and replication-coupled DNA repair. Here we describe how plasmids containing a single ICL can be introduced in the model organism *C. elegans* and their *in vivo* fate be monitored. Transgenes can efficiently be introduced in *C. elegans* through microinjection of DNA into the gonadal syncytium of a young adult hermaphrodite. During oogenesis the exogenously provided DNA fragments will recombine and form arrays, which, at low frequency, can start to behave as a chromosome, *i.e.* be replicated and passed on to progeny cells each cell cycle [33,39]. As a result, a subset of the F1 generation carries such an extrachromosomal array. We mixed the cisplatin crosslink-containing plasmids used in [9] with non-damaged plasmids that encode the mCherry marker and select mCherry positive transgenic F1 animals. In order for the transgenic array to be expressed in multiple cells of the animal it needs to be replicated during embryonic development. The ICLs thus need to be repaired for the host plasmid to become part of a stably transmitted extrachromosomal array. We rationalized that the extrachromosomal arrays in transgenic F1 animals would contain specific footprints of ICL repair, and because dozens of plasmid copies can be enclosed in one array, multiple ICL repair footprints can be obtained in a single transgenic F1. We injected WT animals either with a mix of pICL and mCherry marker plasmids, or with a control mix of pCON (the control plasmid that has the exact same sequence as pICL yet is without the crosslink) and the marker plasmids. All mCherry expressing F1 animals were isolated and allowed to produce offspring before their DNA was extracted. To allow for the analysis of the sequence of the ICL repair products we amplified a locus of approximately 700 nucleotides encompassing the original location of the ICL (Fig. 2b). The sequence at the site of the ICL is constructed in such a way that error-free repair results in the formation of a recognition site for the SapI restriction enzyme (Fig. 2a). This characteristic can be used to estimate the relative level of mutagenic repair at the crosslink versus the level of error-free repair. After PCR amplification the PCR products were incubated with SapI enzyme. When this protocol is performed on transgenic F1's that were injected with pCON control plasmids the PCR product of 727 bp is digested into two fragments of 383 and 344 bp (Fig. 2c, top right). Importantly, this control demonstrates that the SapI digest is almost complete: little or no uncut PCR product is visible on gel. Similar results were obtained for non-transgenic F1's from pCON-injected animals (data not shown);

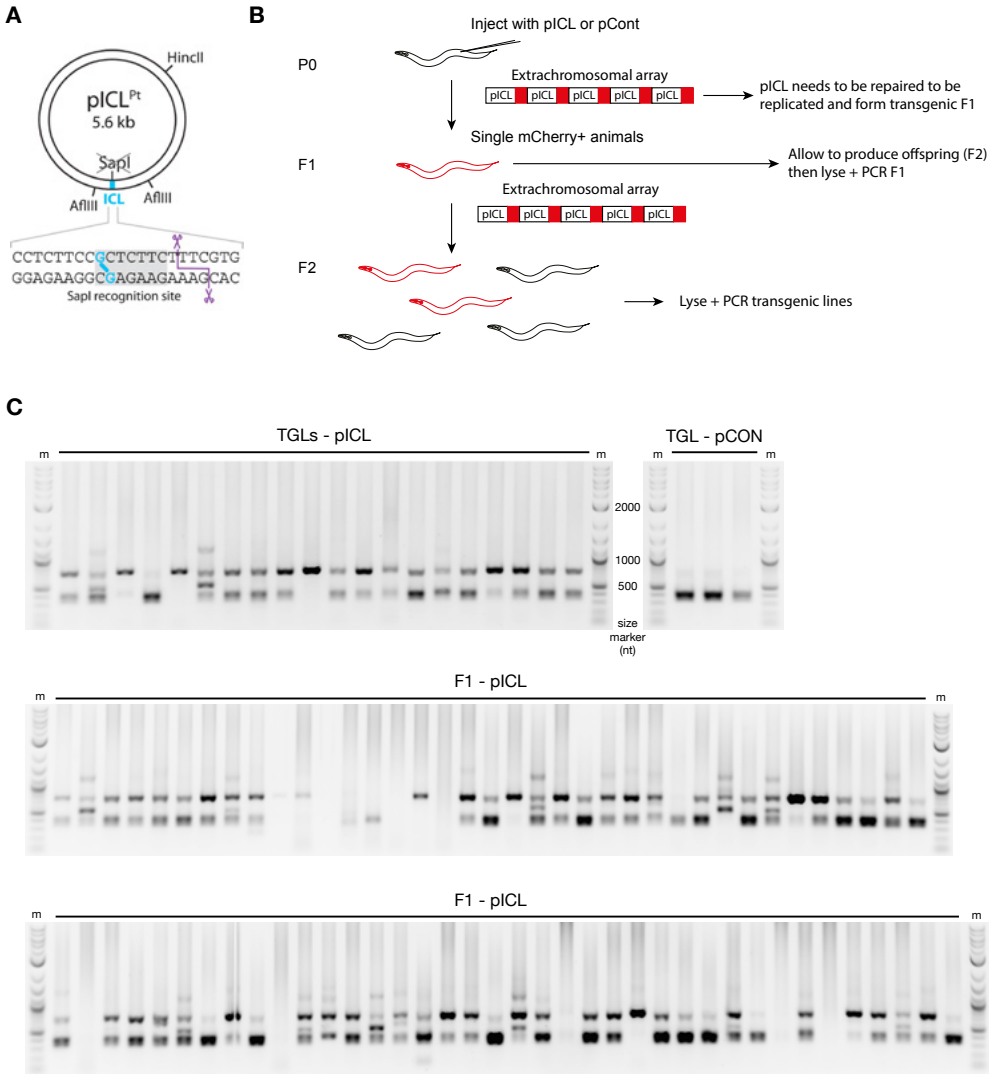


Figure 2. Experimental set-up and proof of principle. *a) Graphic representation of the pICL plasmid used in this study (taken from [10] with permission from AAAS/Science). b) Schematic of the pICL injection protocol in C. elegans. c) Gel electrophoresis images of pICL PCR products treated with SapI restriction enzyme. After injection of WT P0 animals with pICL-mCherry mix or pCON-mCherry mix F1 animals were screened. DNA was extracted from mCherry positive F1's and transgenic lines (TGLs) and used as template. pCON and error-free repaired pICL will result in a PCR product of 727 bp long that is cut into two fragments of 383 and 344 bp by SapI digestion, while mutation at the site of the crosslink result in a SapI resistant fraction of the PCR product. Larger insertions and deletions induce clearly visible changes in the size of PCR products.*

these mCherry expressing animals did not inherit the extrachromosomal array in their germline progenitor cells, and thus didn't pass the array onto subsequent generations.

Error-free repair of the pICL plasmid will also produce a PCR product (727 bp) that is susceptible to SapI digestion, however, a base pair mutation at the site of the crosslink will disrupt the SapI recognition site and prevents SapI cutting, thus resulting a digestion resistant (727 bp) fragment. The same holds for small deletions and insertions that do not induce a recognizable size change of the PCR products. Larger deletions or insertions will result in a PCR fragment of a different size yet also disrupt the SapI recognition sequence. We found that injection of the pICL mix into WT animals produced transgenic progeny indicative of both error-free and error-prone ICL repair. As expected, single transgenic lines (as well as F1's) often show multiple different repair products (Fig. 2c). Our results show that cisplatin ICLs are highly mutagenic, even in a genetic background that is fully proficient in DNA repair.

Characterization of cisplatin ICL repair products in WT animals

Having multiple ICL repair product in a single F1 precludes an immediate quantitative analysis. To solve this caveat and also to study the ICL repair products at nucleotide level we cloned the PCR products and sequenced the clones (see Table 1 and Materials & Methods for experimental details). Taking notice of potential (limited) data skewing because of preferential amplification of smaller than wildtype products, we constructed libraries of ICL repair products from transgenic F1 and transgenic lines. As expected from the data presented in figure 2, we found that in WT animals approximately half of all identified sequences were without a mutation, thus resulting from error-free repair or bypass (Fig. 3a).

		Transgenic F1 animals				Transgenic lines					
Genotype	Injected P0 animals	Transgenic animals singled	Pools prepared for cloning	Clones sequenced[1]	Repair products identified	Transgenic animals singled	Pools prepared for cloning	Clones sequenced[1]	Repair products identified	Total pools	Repair products total[2]
WT	17	107	5	120	106	21	5	120	86	10	192
fcd-2	2	43	2	48	42	4	1	24	10	3	52
xpa-1	5	109	7	168	114	19	5	120	42	12	156
polh-1	8	55	3	72	55	8	2	48	33	5	88
rev-3	8	88	4	96	71	2	TLGs did not produce offspring			4	71
rev-1	30	29	3	72	69	7	No data available			3	69
polk-1	11	25	2	48	42	3	TLGs did not produce offspring			2	42
pcn-1(K165R)	10	88	4	96	76	1	TLGs did not produce offspring			4	76

Table 1. Detailed data on ICL injections. *Detailed data on ICL injections. The table lists the number of injected animals for the indicated genotypes and the resulting F1's, transgenic lines, pools of PCR products, and properly sequenced pICL repair products. [1]Each pool was used for a pGEM-T ligation reaction and E. coli transformation. Of each plate 24 white colonies were picked into ddH2O and this was used as substrate for a colony PCR; then amplification products were sequenced. [2]These numbers are lower than the [total pools] x 24, because not each PCR and/or sequence reaction was successful.*

The mutagenic repair events were given different classifications: single nucleotide variations (SNVs) at or close to the ICL site, inserts without deletions of the original sequence, deletions smaller than 50 bp (with or without insertions), and deletions bigger than 50bp (with or without insertions). SNVs make up ~7% of all repair/bypass products. Insertions without loss of the original sequence are very rare; we only observe a single case in all 192 sequenced repair products. Deletions are much more abundant and are categorized in two distinct size ranges: ~25% of the repair products are small deletions

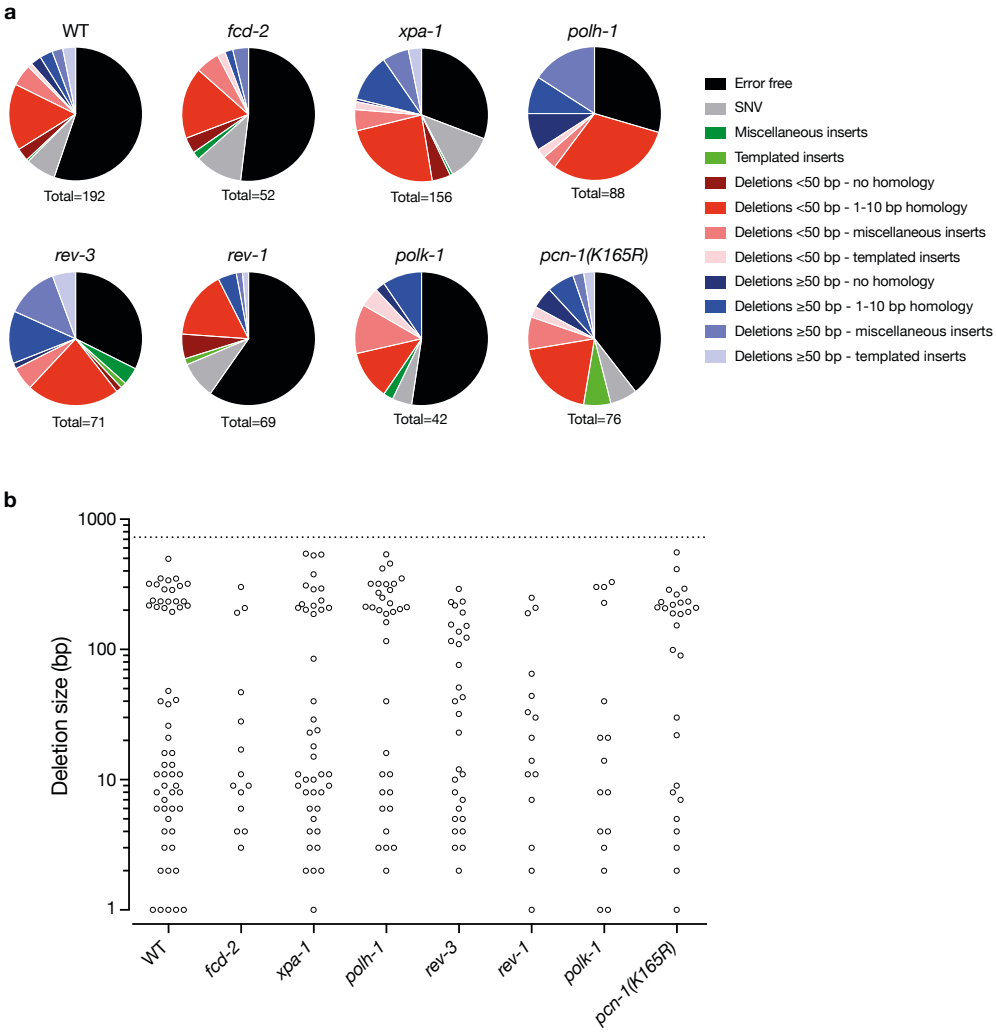


Figure 3. Mutational spectra observed in different DNA response deficient backgrounds. *a) All pICL repair products obtained after sequencing are presented for the indicated genotypes. Categories of repair products were based on observations in WT. b) Sizes of unique deletions observed in the indicated genotypes.*

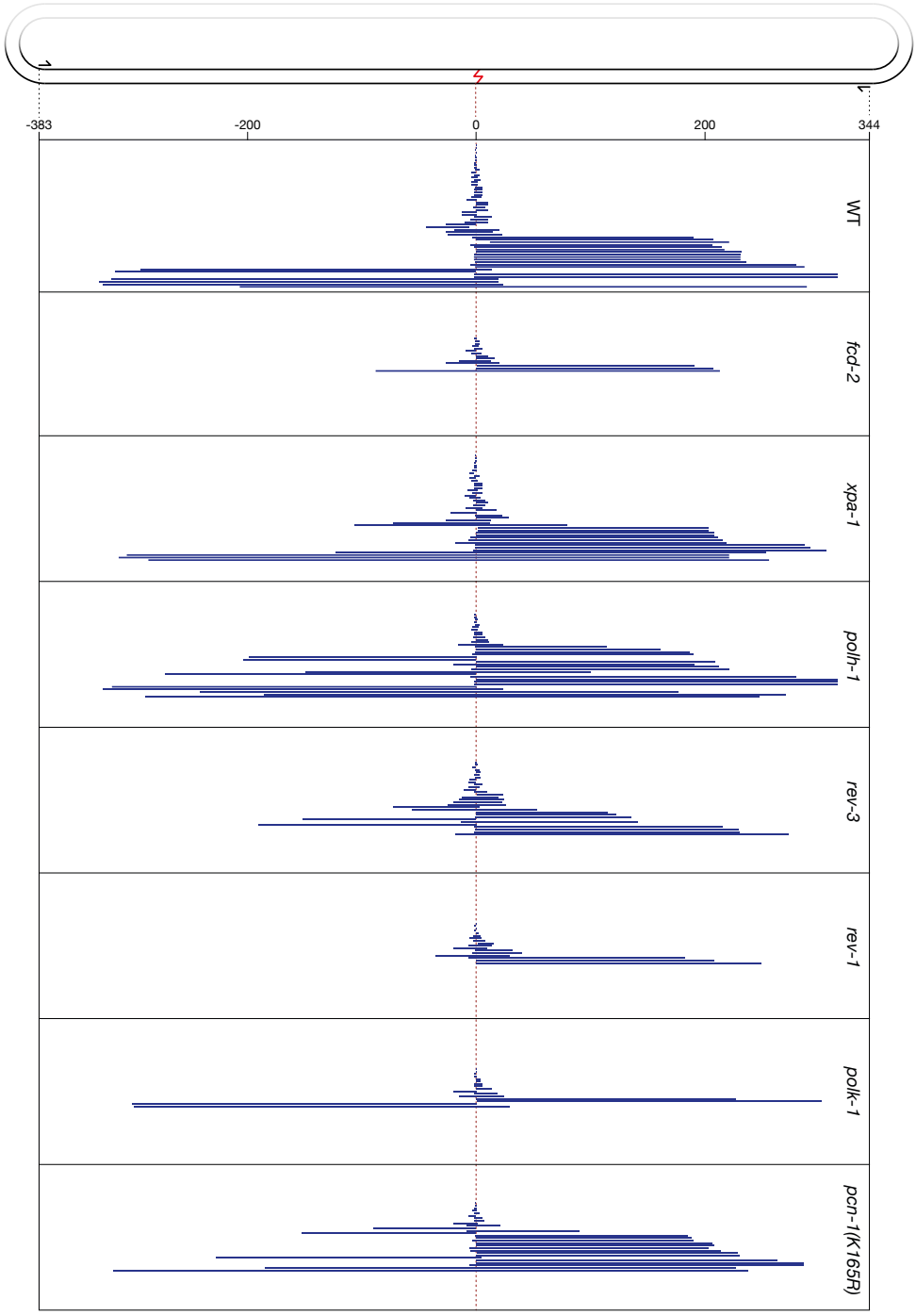
of 1-50 nt and ~11% are deletions >50 nt (Fig. 3 & 4). When studying the 1-50 deletions in further detail we observe characteristic subcategories: deletions where no homology is found at the break ends (~3%), deletions with 1-10 nt homology at the break ends (~16%), deletions with miscellaneous inserts (~3%), and deletions with inserts templated from sequences flanking the deletion (~3%). Within the collection of larger deletions (>50 nt) these same subcategories are found: deletions without homology (~3%), deletions with 1-10 nt homology (~3%), deletions with miscellaneous inserts (~3%), and deletions with inserts templated from sequences flanking the deletion (~3%).

The single nucleotide substitutions are best explained by mutagenic TLS across the unhooked crosslink. A mechanisms that has also been described in Xenopus experiments where in approximately 3% of ICL repair events SNVs are induced by mutagenic TLS [9]. We previously found that persistent replication blocks result in 50-200 bp deletions with significant micro-homology at the deletion junctions together with the occasional presence of templated inserts; a result of polymerase theta mediated end-joining (TMEJ) of replication associated DNA breaks [22,31,40,41]. Also here, we observe deletions of similar size and with other characteristics of TMEJ. This outcome could result from the inability to TLS past the unhooked crosslink. Interestingly, we also observed a high frequency of deletions that are smaller than 50 bp in size - a size category that never substantially contributes to the deletions resulting from persistent replication blocks present in one DNA strand (G-quadruplex structures, mono-adducts and UV-lesions), suggesting that these small deletions are specific for ICLs, a notion that is supported by the observation of similar deletion sizes in *C. elegans* exposed to the ICL-inducing agents mechlorethamine, cisplatin and MMC [27,32]. Interestingly, these smaller deletions also show characteristics of TMEJ. We recently found that one of the two junctions of deletions resulting from persistent replication blocks is exactly at the position of the replication block, best illustrated by deletions at G-quadruplex sites [40,41]. We here find that the vast majority of deletions originating from ICLs also have one junction immediately flanking the lesion (Figure 4), which may point towards a repair intermediate where one of the crosslinked bases block polymerase action.

Multiple DNA damage response pathways are involved in cisplatin ICL repair

The wide spectrum of repair outcomes found in WT animals suggests the involvement of multiple DNA repair pathways. We thus tested a set of genetic backgrounds deficient in different aspects of the DNA damage response: TLS *(polh-1, polk-1, rev-1, rev-3* and *pcn-1(K165R))*; NER *(xpa-1)*; and ICL repair (*fcd-2*). Experimental details for the different

Figure 4. Unique deletions in relation to the location of the crosslink. *On the top a schematic representation of the pICL plasmid with primer sites indicated by small arrows. Unique deletions isolated from the indicated genotypes are mapped to the pICL sequence. Each bar represents a deletion and the red dotted line indicates the location of the crosslink.*


1
2
-383
-200
0
200
344
WT
fcd-2
xpa-1
polh-1
rev-3
rev-1
polk-1
pcn-1(K165R)

strains can be found in Table 1 and in the section Material & Methods. All the factors analyzed did not affect transgenic array formation in the absence of the ICL lesion (Robin van Schendel, M.T. unpublished observation).

We observed a significant shift towards mutagenic repair in strains that lack the TLS polymerases Polζ *(rev-3*) and Polη (*polh-1*) and in animals deficient for the NER protein XPA-1 (Table 2), arguing that both NER and TLS acts to repair ICL in this context. Together with the broad spectrum of outcomes in wild type, this result validates the assay we here present as a means to study ICL repair *in vivo* at the nucleotide level. Below, we will briefly discuss the preliminary data derived in the different genetic backgrounds. While these results are preliminary, they can serve as a platform to guide future research.

FANCD2

Although *C. elegans fcd-2* mutants are sensitive to ICL-inducing agents [25,31], the loss of this ICL repair factor does not have a significant effect on the mutagenicity and the kind of mutations that occur at a well-defined ICL in exogenously provided plasmids (Table 2, Fig. 3 & 4). At present we do not have an explanation for this outcome apart from suggesting that the type of ICL-repair we assay is independent of the Fanconi anemia pathway.

XPA

We find that deficiency for the NER factor XPA (*xpa-1*) results in a statistically significant shift towards more mutagenic repair in comparison with error-free repair (Table 2). Especially deletion frequencies are increased (Fig. 3 & 4). These findings support a role of XPA in error-free repair of ICL repair in *C. elegans*. One explanation for our observation is that loss of or incomplete replication-independent ICL repair leads to more replication blocks that subsequently cause the formation of DNA breaks and deletions. Although there are many indications that multiple NER proteins are involved in ICL repair, their role in this pathway is not fully understood [17]. XPA probably plays a role in damage recognition or unwinding of the DNA around the ICL because it binds junctions between single- and double stranded DNA [15], and it seems to do this especially in replication-independent ICL repair [16]. In addition, NER may play a role in the removal of the unhooked crosslink that remains after TLS has progressed.

Translesion synthesis

We did not observe any effect when *polk-1* was mutated (Table 2, Fig, 3 & 4), as could be expected from the results of the cisplatin and UV-TMP sensitivity experiments presented in Figure 1. Interestingly, the ICL repair spectrum derived from *rev-1* mutant animals is also indistinguishable from that derived from wild type animals (Table 2, Fig, 3 & 4), while these mutant animals are highly sensitive to ICL-inducing

	Repair products		Chi-square test	
Genotype	Error-free (n)	Mutagenic (n)	p-values	Significantly diffent from WT (p<0.0014)
WT	106	86	-	-
fcd-2	27	25	0,633790794	FALSE
xpa-1	48	108	8,34203E-10	TRUE
polh-1	26	62	1,2911E-06	TRUE
rev-3	23	48	0,000110767	TRUE
rev-1	40	27	0,459551655	FALSE
polk-1	22	20	0,712519073	FALSE
pcn-1(K165R)	30	46	0,005807866	FALSE

Table 2. Statistical analyses of error-free and mutagenic repair in different genotypes as compared to WT. *The numbers of error-free repair vs. mutagenic repair products in all genetic backgrounds were used to determine statistic differences of the indicated genetic backgrounds. P-values were calculated with Chi-square tests. In order to control for repeated testing Bonferroni correction was performed on the original α=0.01: seven different genotypes were each compared to WT leading to a new α=0,01/7=0.0014. A genotype is significantly different from WT when p<α.*

agents (Fig. 1). It may be that REV-1 is especially important for TLS of mono-adducts that are induced upon exposure to these agents. It is also possible that one of the other TLS polymerases can take over, resulting in less efficient TLS and high cytotoxicity, but no substantial changes in mutation induction. Our *in vivo* findings are different from observations in Xenopus egg extracts where it was shown that the induction of SNVs close to the ICL is dependent on REV1 [12]. Animals that contain the *pcn-1(K165R)* allele had a similar mutation profile as WT animals (Table 2, Fig. 3 & 4), while being more sensitive to ICL-inducing agents. Mono-ubiquitination of PCNA is a general activation signal for TLS and therefore the effect may be similar as loss of REV-1, and specifically acting on mono-adducts or reduced TLS (causing cytotoxicity), but no change in mutagenicity.

The most striking changes are observed in *polh-1* and *rev-3* mutant animals (Table 2, Fig. 3 & 4). A complete lack of SNVs and a substantial increased large deletion frequency is observed in both these mutants. Thus, contrary to REV1, a role of Polζ appears to be conserved between *C. elegans* and Xenopus [12]. Our findings suggest a role for Polζ and Polη TLS polymerases in the mutagenic bypass of unhooked crosslinks. Our previous research demonstrated that loss of TLS causes persistent replications blocks that in turn lead to DSBs. Repair of these breaks produces deletions with a TMEJ footprint and this is also what we observe here [22,23]. Specifically, the increased frequency of 50-200 deletions is in line with this notion. In such a scenario, replication at the unhooked crosslink is completely blocked ultimately leading to deletion formation at such sites.

Conclusions & Perspectives

ICL-inducing agents are the most widely-used class of chemotherapeutic. They are so effective because DNA crosslinks are specifically hard to tolerate in rapidly cycling cells. However, despite intense research efforts over the last two decades, how cells act to repair these harmful lesions is still incompletely understood. Studying the cellular responses to ICLs may help to develop novel and improved chemotherapeutics. In addition, cancer cells can adopt to the presence of ICLs, for example by upregulating DNA damage response pathways leading to chemoresistance, a major hurdle in the treatment of cancer. Previous work has firmly established a role for TLS, FA and HR pathways working together in ICL repair, thereby contributing to DNA damage tolerance [Reviewed in 42]. With the *in vivo* assay we present here, we wished to expand the toolbox for research into the molecular repair processes of DNA interstrand crosslinks, and by comparing the outcomes to NGS data of animals exposed to ICL-inducing agents (R. van Schendel and M.T. unpublished results) provide more clarity into substrate specificity of different repair mechanisms.

The here-described assay still requires further development. For instance, to monitor reduced array formation we can co-inject an undamaged plasmid that is identical to the ICL-containing plasmid yet carrying a limited number of SNVs, as well as titrating the number of repair events to less than one per animal - at this moment we fail to monitor drop-out effects where mutant animals may produce less transgenic F1's, or F1's carry relatively low number of copies of the pICL plasmid in the transgenic array. Another valuable development is NGS of PCR products of large pools of animals to provide stronger statistical power.

Future and ongoing work is directed to expand on the analysis here presented, also studying plasmids that carry a psoralen crosslink [13]. Taking the latest advances in genome editing in consideration it may become possible to insert small pieces of DNA that carry a crosslink or other lesion of choice directly into the genome as has recently been done in *E. coli* [43].

In conclusion, we here present a novel assay to study the repair and mutagenesis of a single crosslinks at a known site in the nematode *C. elegans* in which many ICL repair factors are conserved [24,25]. With this assay the *in vitro* studies performed in Xenopus egg extract which have provided a detailed model for ICL repair can be paralleled, and other questions specific to an *in vivo* context can be addressed.

Acknowledgements

Several strains were provided by the CGC, which is funded by the NIH Office of Research Infrastructure Programs . This work is supported by the European Research Council (ERC) and the Netherlands Organization for Scientific Research (NWO). We are grateful to Dr. Puck Knipscheer for providing us with the pICL and pCON plasmids.

References

[1] Deans AJ, West SC. DNA interstrand crosslink repair and cancer. Nat Rev Cancer 2011;11:467–80. doi:10.1038/nrc3088.

[2] Scharer OD. DNA Interstrand Crosslinks: Natural and Drug-Induced DNA Adducts that Induce Unique Cellular Responses. ChemBioChem 2005.

[3] Goodman LS, Wintrobe MM, Dameshek W, Goodman MJ, GILMAN A, McLennan MT. Nitrogen mustard therapy: Use of Methyl-Bis(Beta-Chloroethyl) amine Hydrochloride and Tris(Beta-Chloroethyl) amine Hydrochloride for Hodgkin's Disease, Lymphosarcoma, Leukemia and Certain Allied and Miscellaneous Disorders. Journal of the American Medical Association 1946;132:126–32. doi:10.1001/jama.1946.02870380008004.

[4] Nalepa G, Clapp DW. Fanconi anaemia and cancer: an intricate relationship. Nat Rev Cancer 2018;18:168–85. doi:10.1038/nrc.2017.116.

[5] Niedernhofer LJ, Lalai AS, Hoeijmakers JHJ. Fanconi anemia (cross)linked to DNA repair. Cell 2005;123:1191–8. doi:10.1016/j.cell.2005.12.009.

[6] Ceccaldi R, Sarangi P, D'Andrea AD. The Fanconi anaemia pathway: new players and new functions. Nature Reviews Molecular Cell Biology 2016;17:337–49. doi:10.1038/nrm.2016.48.

[7] Walden H, Deans AJ. The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. Annu Rev Biophys 2014;43:257–78. doi:10.1146/annurev-biophys-051013-022737.

[8] Muniandy PA, Liu J, Majumdar A, Liu S-T, Seidman MM. DNA interstrand crosslink repair in mammalian cells: step by step. Crit Rev Biochem Mol Biol 2010;45:23–49. doi:10.3109/10409230903501819.

[9] Räschle M, Knipscheer P, Knipsheer P, Enoiu M, Angelov T, Sun J, et al. Mechanism of replication-coupled DNA interstrand crosslink repair. Cell 2008;134:969–80. doi:10.1016/j.cell.2008.08.030.

[10] Knipscheer P, Räschle M, Smogorzewska A, Enoiu M, Ho TV, Schärer OD, et al. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science 2009;326:1698–701. doi:10.1126/science.1182372.

[11] Long DT, Räschle M, Joukov V, Walter JC. Mechanism of RAD51-dependent DNA interstrand cross-link repair. Science 2011;333:84–7. doi:10.1126/science.1204258.

[12] Budzowska M, Graham TGW, Sobeck A, Waga S, Walter JC. Regulation of the Rev1-pol ζ complex during bypass of a DNA interstrand cross-link. Embo J 2015;34:1971–85. doi:10.15252/embj.201490878.

[13] Semlow DR, Zhang J, Budzowska M, Drohat AC, Walter JC. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase. Cell 2016;167:498–511.e14. doi:10.1016/j.cell.2016.09.008.

[14] Williams HL, Gottesman ME, Gautier J. The differences between ICL repair during and outside of S phase. Trends in Biochemical Sciences 2013;38:386–93. doi:10.1016/j.tibs.2013.05.004.

[15] Camenisch U, Dip R, Schumacher SB, Schuler B, Naegeli H. Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. Nat Struct Mol Biol 2006;13:278–84. doi:10.1038/nsmb1061.

[16] Shen X, Do H, Li Y, Chung W-H, Tomasz M, de Winter JP, et al. Recruitment of fanconi anemia and breast cancer proteins to DNA damage sites is differentially governed by replication. Mol Cell 2009;35:716–23. doi:10.1016/j.molcel.2009.06.034.

[17] Wood RD. Mammalian nucleotide excision repair proteins and interstrand crosslink repair. Environ Mol Mutagen 2010;51:520–6. doi:10.1002/em.20569.

[18] Sharma S, Canman CE. REV1 and DNA polymerase zeta in DNA interstrand crosslink repair. Environ Mol Mutagen 2012;53:725–40. doi:10.1002/em.21736.

[19] Roy U, Schärer OD. Involvement of translesion synthesis DNA polymerases in DNA interstrand crosslink repair. DNA Repair 2016;44:33–41. doi:10.1016/j.dnarep.2016.05.004.

[20] Williams HL, Gottesman ME, Gautier J. Replication-independent repair of DNA interstrand crosslinks. Mol Cell 2012;47:140–7. doi:10.1016/j.molcel.2012.05.001.

[21] Roerink SF, Koole W, Stapel LC, Romeijn RJ, Tijsterman M. A Broad Requirement for TLS Polymerases η and κ, and Interacting Sumoylation and Nuclear Pore Proteins, in Lesion Bypass during *C. elegans* Embryogenesis. PLoS Genet 2012;8:e1002800. doi:10.1371/journal.pgen.1002800.

[22] Roerink SF, van Schendel R, Tijsterman M. Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. Genome Research 2014;24:954–62. doi:10.1101/gr.170431.113.

[23] van Bostelen I, Tijsterman M. Combined loss of three DNA damage response pathways renders *C. elegans* intolerant to light. DNA Repair 2017. doi:10.1371/journal.pbio.1001378.

[24] Youds JL, Barber LJ, Boulton SJ. *C. elegans*: a model of Fanconi anemia and ICL repair. Mutat Res 2009;668:103–16. doi:10.1016/j.mrfmmm.2008.11.007.

[25] Collis SJ, Barber LJ, Ward JD, Martin JS, Boulton SJ. *C. elegans* FANCD2 responds to replication stress and functions in interstrand cross-link repair. DNA Repair 2006;5:1398–406. doi:10.1016/j.dnarep.2006.06.010.

[26] Povirk LF, Shuker DE. DNA damage and mutagenesis induced by nitrogen mustards. Mutat Res 1994;318:205–26.

[27] Meier B, Cooke SL, Weiss J, Bailly AP, Alexandrov LB, Marshall J, et al. *C. elegans*whole-genome sequencing reveals mutational signatures related to carcinogens and DNA repair deficiency. Genome Research 2014;24:1624–36. doi:10.1101/gr.175547.114.

[28] Lemaire MA, Schwartz A, Rahmouni AR, Leng M. Interstrand cross-links are preferentially formed at the d(GC) sites in the reaction between cis-diamminedichloroplatinum (II) and DNA. Proc Natl Acad Sci U S A 1991;88:1982–5. doi:10.1073/pnas.1718185115.

[29] Yandell MD, Edgar LG, Wood WB. Trimethylpsoralen induces small deletion mutations in Caenorhabditis elegans. Proc Natl Acad Sci U S A 1994;91:1381–5. doi:10.1073/pnas.91.4.1381.

[30] Barstead RJ, Moerman DG. *C. elegans* deletion mutant screening. Methods Mol Biol 2006;351:51–8. doi:10.1385/1-59745-151-7:51.

[31] van Schendel R, van Heteren J, Welten R, Tijsterman M. Genomic Scars Generated by Polymerase Theta Reveal the Versatile Mechanism of Alternative End-Joining. PLoS Genet 2016;12:e1006368. doi:10.1371/journal.pgen.1006368.

[32] Tam AS, Chu JSC, Rose AM. Genome-Wide Mutational Signature of the Chemotherapeutic Agent Mitomycin C in Caenorhabditis elegans. G3 (Bethesda) 2015;6:133–40. doi:10.1534/g3.115.021915.

[33] Stinchcomb DT, Shaw JE, Carr SH, Hirsh D. Extrachromosomal DNA transformation of Caenorhabditis elegans. Mol Cell Biol 1985;5:3484–96.

[34] Brenner S. The genetics of Caenorhabditis elegans. Genetics 1974;77:71–94.

[35] Frøkjær-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen S-P, et al. Single-copy insertion of transgenes in Caenorhabditis elegans. Nat Genet 2008;40:1375–83. doi:10.1038/ng.248.
[36] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–10. doi:10.1016/S0022-2836(05)80360-2.
[37] Eastman A. Reevaluation of interaction of cis-dichloro (ethylenediamine) platinum (II) with DNA. Biochemistry 1986. doi:10.1021/bi00361a026.
[38] Dronkert ML, Kanaar R. Repair of DNA interstrand cross-links. Mutat Res 2001;486:217–47. doi:10.1016/S0921-8777(01)00092-1.
[39] Mello CC, Kramer JM, Stinchcomb D, Ambros V. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. Embo J 1991;10:3959–70.
[40] Koole W, van Schendel R, Karambelas AE, van Heteren JT, Okihara KL, Tijsterman M. A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat Comms 2014;5:3216. doi:10.1038/ncomms4216.
[41] Lemmens B, van Schendel R, Tijsterman M. Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. Nat Comms 2015;6:8909. doi:10.1038/ncomms9909.
[42] Haynes B, Saadat N, Myung B, Shekhar MPV. Crosstalk between translesion synthesis, Fanconi anemia network, and homologous recombination repair pathways in interstrand DNA crosslink repair and development of chemoresistance. Mutat Res Rev Mutat Res 2015;763:258–66. doi:10.1016/j.mrrev.2014.11.005.
[43] Pagès V, Fuchs RP. Inserting Site-Specific DNA Lesions into Whole Genomes. Methods Mol Biol 2018;1672:107–18. doi:10.1007/978-1-4939-7306-4_9.