Alternative end-joining of DNA breaks

Schendel, Robin van

Citation

Schendel, R. van. (2016, December 15). *Alternative end-joining of DNA breaks*. Retrieved from https://hdl.handle.net/1887/45030

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/45030

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/45030 holds various files of this Leiden University dissertation

Author: Schendel, Robin van
Title: Alternative end-joining of DNA breaks
Issue Date: 2016-12-15

GENERAL DISCUSSION AND FUTURE PERSPECTIVES

6

This dissertation discusses several aspects of double-strand break (DSB) repair in *C. elegans*. Faithfull repair of DSBs is crucial for cells to maintain genome stability and for that reason eukaryotic cells are equipped with a variety of DSB-repair mechanisms. Apart from homologous recombination (HR), which is considered to be an error-free pathway, most of the other DSB-repair pathways are intrinsically error-prone, frequently leading to small genetic changes, but occasionally leading to gross chromosomal aberrations. Cells that are compromised in their ability to repair DSBs are more likely to undergo malignant transformation. Although all cells within a single organism are generally equipped with the same mechanisms to repair DSBs, the contribution and availability of each repair mechanism depends on cell type (e.g. germ cells versus somatic cells) and cell stage. It is especially crucial for germ and stem cells to properly deal with genetic insults as these cells give rise to progenitors.

To investigate DSB-repair pathways I made use of whole-genome sequencing approaches, which enabled me to examine the entire genome of animals that were either wild type or carried a genetic defect in one or more DNA repair mechanisms. By probing the genomes of animals that accumulated mutations we identified specific signatures, one of which leading to the identification of a previously unknown error-prone DSB-repair pathway, which depends on the A-family polymerase Theta (POLQ). In essence, POLQ attempts to connect two DNA ends by using single base-pair of homology between the ends from which POLQ can extend. This frequently results in the repair of the break and the deletion of a small piece of genetic information. Occasionally, however, during extension the two DNA ends dissociate and the process of connecting and extension by POLQ is repeated until the break is repaired. The repeated action of POLQ leaves behind a smoking-gun for POLQ-mediated repair: a small piece of newly synthesized DNA, which is a carbon copy of part of the DSB flank is inserted between the two broken ends.

Although this thesis provides detailed mechanistic insight into how POLQ-mediated end-joining repairs a break *in vivo*, many questions still remain unanswered. Especially little is currently known about the spatial and temporal regulation of this pathway as well as the context in which this pathway operates. A selection of outstanding questions will be discussed in the following sections.

How is Polymerase Theta-mediated repair orchestrated?

Our laboratory has shown that POLQ plays an important role in maintaining genome stability, but it remains unknown how POLQ is recruited to sites of damage. The primary DNA-damage sensors ATM and ATR are conserved in *C. elegans* but it is currently an outstanding question whether the downstream targets of these signalling kinases are conserved as well. Although both ATM and ATR share many downstream targets, at least in higher eukaryotes, they respond to different types of damages. ATR primarily responds to stalled replication forks lesions, while ATM is activated by DSBs[1]. Mice with defects in both ATM and POLQ exhibit a more severe phenotype than either deficiency alone, suggesting that POLQ and ATM do not act in the same pathway[2]. Unfortunately, *C. elegans* ATR is an essential protein making it impossible to genetically address its involvement in POLQ-mediated repair.

Some of our data indicate that POLQ acts at replication-associated DSBs (Chapter 3 and [3]) as the absence of TLS polymerases pol eta and kappa as well as the helicase FANCJ result in a distinct class of deletions that for their formation depend on POLQ. One possibility would therefore be that POLQ is recruited to DSBs by factors involved in replication. However, both transposon and

CRISPR\Cas9-induced breaks, which are thought to form independent of replication also require POLQ activity for their repair (Chapter 4). This suggests that POLQ can be recruited to DSB outside the context of DNA replication. A candidate for this function is the ssDNA binding protein RPA that coats ssDNA of resected DSBs to protect it from degradation. Notably, RPA coats ssDNA in- and outside the context of replication, which fits with POLQ recruitment for replication-associated and replication-uncoupled breaks.

Some proteins have already been implicated in alternative end-joining (altEJ), an ill-defined category of DSB-repair pathways that includes POLQ-mediated repair. PARP1 is implicated in altEJ and is rapidly recruited to DSBs. In mammalian systems PARP1 was shown to act upstream of POLQ, though it is yet unclear whether PARP1 recruits POLQ directly or indirectly and in which context[4,5]. Surprisingly, our preliminary data in *C. elegans* suggest that animals deficient for *parp-1* do not show a DSB-repair defect and are still proficient in POLQ-mediated repair, arguing that in *C. elegans*, POLQ action does not depend on PARP.

A second question that is currently unanswered is which factor(s) are responsible for the finalization of repair in POLQ-mediated repair? The current model for POLQ-mediated repair requires a ligation step. A likely candidate is LIG3, also because this protein has previously been implicated in altEJ. In mice, $LIG3^{-/-}$ cells could be created but only when LIGI or LIG3 was targeted to the mitochondria. It was subsequently found that the frequency of altEJ-mediated DNA translocations was reduced in a nuclear LIG3-deficient mouse backgrounds when breaks were induced by a zinc-finger endonuclease, implicating LIG3 in altEJ[6]. No mutant allele of *C. elegans* LIG-3 (K07C5.3) is currently available, but I have recently used CRISPR\Cas9-induced mutagenesis to create one, which is currently being investigated for POLQ-mediated repair phenotypes.

Instead of using a candidate approach to identify factors that are involved in POLQ-mediated repair we can perform unbiased screens. The classical approach in *C. elegans* is to carry out a forward genetic screen combined with a phenotypical read out to identify mutants of interest. A pilot EMS screen was performed that identified two new alleles of POLQ but thus far no novel factors. Because this was a very small-scale being far from saturated I suggest increasing scale.

An alternative approach is to use a biochemical approach: immunoprecipitation (IP) of POLQ followed by mass spectrometry to identify proteins that co-precipitate, indicating a direct or indirect interaction with POLQ. For years it has been technically extremely challenging to endogenously tag proteins in *C. elegans*, but CRISPR\Cas9 technology made it feasible to tag proteins with for example GFP or FLAG, thus enabling us to IP POLQ. The latter approach would also allow the identification of essential genes that would be missed in forward genetics screens.

Which parameters determine the deletion size in Polymerase Theta-mediated repair?

One of the most enigmatic questions that thus far remains unanswered is what determines the deletion size in POLQ-mediated repair events? The heritable genomic changes seen after repair of transposition and CRISPR\Cas-9 breaks are typically <20 bp, while for replication-associated deletions they are 50 – 300 bp, sometimes larger, but almost never smaller. Can the difference between repair outcomes of direct breaks (e.g. via CRISPR\Cas-9 or transposition) and replication-associated breaks simply be explained by the context in which the break occurs? Moreover, we found subtle but clear differences between the deletion-size distribution of TLS-deficient and FANCJ-deficient animals: intriguingly, when we compare both distributions we find a median

deletion size of 110 and 138 bp respectively (*n* > 90 for both sets)[3,7]. This difference is most probably explained by the fact that a G-quadruplex motif, being 20-25 bases on average, is the replication-blocking obstacle in FANCJ-deficient animals, while a single damaged base blocks the fork in TLS-deficient animals. This notion argues that the context of the replication fork impediment is of direct influence to the resulting genomic change and is thus a factor of relevance in thinking about the mechanism.

At present, we do not know whether a G4-structure is more likely to occur in the leading or lagging strand. Our data demonstrates that replication can approach a replication block (e.g. a G4-structure or a psoralen cross-link) to within a few nucleotides ([3], Chapter 3 and 5), and as such determines one deletion breakpoint. But what determines the other breakpoint, and thus the size of the deletion? If the lesion is present in the lagging strand, the other breakpoint may be determined by the previous Okazaki fragment. Okazaki fragments are deposited at ~300 bp intervals[8] which would fit the ~50-300 bp deletion size distribution, about half an Okazaki fragment. On the other hand, if the replication block occurs on the leading strand we foresee two options that can lead to a deletion: re-priming of the leading strand behind the replication blocking lesion or the approach of a converging replication fork. PrimPol, a protein that contains both TLS and primase activity, has been shown to be able to bypass replication blocking lesions either by employing its TLS activity or by re-priming downstream of the blocking lesion[9]. *C. elegans* does not contain a homolog of human PrimPol which makes a jump-over model by re-priming the leading strand downstream of the replication block less likely. In a converging replication fork model a ssDNA gap results of a size that is dependent on how close a converging fork can approach an arrested fork. In the next cell cycle such a ssDNA gap can be converted into a DSB. We have recently provided strong experimental evidence for this scenario[10]. To demonstrate that Okazaki fragments are of relevance in deletion formation we need to perturb Okazaki fragment deposition. To address the question whether G4 structures are predominantly forming in leading or lagging strands, we require information on origins of replication.

Is the role of Polymerase Theta conserved in higher eukaryotes?

To understand DSB-repair in model organisms such as *C. elegans* is not our primary goal. POLQ is conserved in mouse and human, but only recently it became evident that the role of POLQ in DSB-repair is also functionally conserved[4,5,11-13]. It is thus of great interest to translate the findings observed in model systems to humans. Our laboratory found that most deletions that occur in *C. elegans* germ cells are brought about by the activity of POLQ. Sequencing of natural isolates of *C. elegans* have allowed us to examine genome diversification and to discover that genomic changes >1 bp are carrying the hallmarks of POLQ-mediated repair. This specific mutation profile was recapitulated in a small-scale evolution experiment where POLQ-deficient and proficient animals were grown in parallel for ~250 generations. The mutational spectrum observed in POLQ-proficient animals was nearly identical to the spectra observed in natural isolates, but was, however, completely altered in POLQ-deficient animals. From this we concluded that POLQ plays a major role in the genome diversification of *C. elegans*. It will now be of interest to address the contribution of POLQ-mediated repair in genome variations in mammals, either in germ cells leading to genetic variation or in somatic cells leading to cancer.

It is currently unknown why NHEJ does not act on breaks in *C. elegans* germ cells, while it appears to be functional in these cells[14,15]. Both studies show that NHEJ is actively suppressed to

prevent illegitimate repair between chromosomes during meiosis. Given that error-prone repair in germ cells of *C. elegans* almost exclusively rely on POLQ for the repair of DSBs it is of great interest to investigate whether germ cells of higher eukaryotes equally depend on POLQ. Several studies have already identified altered expression profiles for key DSB-repair proteins in germ cells of mice as well as germline mutations that hint towards repair activity by POLQ[16-18]. In the soma the situation is quite different as NHEJ is the dominant pathway to repair spontaneous DSBs that are replication-uncoupled, both in mammals and *C. elegans*[19,20]. It appears that in this context repair by POLQ is rather an alternative to NHEJ and HR as POLQ events can generally only be detected in the absence of one of these DSB-repair pathways. Interestingly, tumours that are HR-deficient rely on POLQ for their survival and knockdown of POLQ in HR-proficient cells upregulates HR activity indicating that they can act on similar substrates. POLQ is therefore considered to be an attractive novel druggable candidate target for cancer therapy[5].

In Chapter 5 we described the *in vivo* mechanism and identified several hallmarks of POLQ-mediated DSB repair. Especially templated flank insertions, where a small piece of DNA identical to nearby sequences is found inserted into a candidate DSB site is a smoking-gun for POLQ-mediated repair. It will therefore be of great interest to query datasets (*e.g.* human tumour datasets and/or *de novo* mutations) for POLQ signatures. A number of reports already anecdotally describe the presence of small insertions that resemble the immediate flank[21-23]. Human dataset generally consists of a mixture of mutational signatures generated by several repair pathways[24]. Dissecting the contribution of each mutational process, including POLQ-mediated repair, will be an interesting challenge for the years to come.

REFERENCES

1 A. M. Weber and A. J. Ryan ATM and ATR as therapeutic targets in cancer *Pharmacol. Ther.* **149**, 124 (2015).
2 N. Shima, R. J. Munroe, and J. C. Schimenti The mouse genomic instability mutation chaos1 is an allele of Polq that exhibits genetic interaction with Atm *Mol. Cell Biol.* **24**(23), 10381 (2004).
3 W. Koole, *et al.* A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites *Nat. Commun.* **5**, 3216 (2014).
4 P. A. Mateos-Gomez, *et al.* Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination *Nature* **518**(7538), 254 (2015).
5 R. Ceccaldi, *et al.* Homologous-recombination-deficient tumours are dependent on Poltheta-mediated repair *Nature* **518**(7538), 258 (2015).
6 D. Simsek, *et al.* DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation *PLoS. Genet.* **7**(6), e1002080 (2011).
7 S. F. Roerink, Schendel R. van, and M. Tijsterman Polymerase theta-mediated end joining of replication-associated DNA breaks in C. elegans *Genome Res.* **24**(6), 954 (2014).
8 G. Abdurashidova, *et al.* Start sites of bidirectional DNA synthesis at the human lamin B2 origin *Science* **287**(5460), 2023 (2000).
9 S. Garcia-Gomez, *et al.* PrimPol, an archaic primase/polymerase operating in human cells *Mol. Cell* **52**(4), 541 (2013).
10 B. Lemmens, Schendel R. van, and M. Tijsterman Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers *Nat. Commun.* **6**, 8909 (2015).
11 T. Kent, *et al.* Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase theta *Nat. Struct. Mol. Biol.* **22**(3), 230 (2015).
12 K. E. Zahn, *et al.* Human DNA polymerase theta grasps the primer terminus to mediate DNA repair *Nat. Struct. Mol. Biol.* **22**(4), 304 (2015).
13 M. J. Yousefzadeh, *et al.* Mechanism of suppression of chromosomal instability by DNA polymerase POLQ *PLoS. Genet.* **10**(10), e1004654 (2014).
14 A. Adamo, *et al.* Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia *Mol. Cell* **39**(1), 25 (2010).
15 B. B. Lemmens, N. M. Johnson, and M. Tijsterman COM-1 promotes homologous recombination during Caenorhabditis elegans meiosis by antagonizing Ku-mediated non-homologous end joining *PLoS. Genet.* **9**(2), e1003276 (2013).
16 C. M. Carvalho, *et al.* Replicative mechanisms for CNV formation are error prone *Nat. Genet.* **45**(11), 1319 (2013).
17 G. Hamer, *et al.* Function of DNA-protein kinase catalytic subunit during the early meiotic prophase without Ku70 and Ku86 *Biol. Reprod.* **68**(3), 717 (2003).
18 M. J. Ashwood-Smith and R. G. Edwards DNA repair by oocytes *Mol. Hum. Reprod.* **2**(1), 46 (1996).
19 A. J. Hartlerode and R. Scully Mechanisms of double-strand break repair in somatic mammalian cells *Biochem. J.* **423**(2), 157 (2009).
20 D. B. Pontier and M. Tijsterman A robust network of double-strand break repair pathways governs genome integrity during C. elegans development *Curr. Biol.* **19**(16), 1384 (2009).
21 J. M. Kidd, *et al.* A human genome structural variation sequencing resource reveals insights into mutational mechanisms *Cell* **143**(5), 837 (2010).
22 D. F. Conrad, *et al.* Mutation spectrum revealed by breakpoint sequencing of human germline CNVs *Nat. Genet.* **42**(5), 385 (2010).
23 C. M. Carvalho, *et al.* Replicative mechanisms for CNV formation are error prone *Nat. Genet.* **45**(11), 1319 (2013).
24 T. Helleday, S. Eshtad, and S. Nik-Zainal Mechanisms underlying mutational signatures in human cancers *Nat. Rev. Genet.* **15**(9), 585 (2014).