DNA-RNA hybrids: the risks of DNA breakage during transcription

Andrés Aguilera & Belén Gómez-González
Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER,
Universidad de Sevilla-CSIC-Universidad Pablo de Olavide,
Seville, Spain
aguilo@us.es

ABSTRACT

Although R loops can occur at different genomic regions, a yet unsolved question is how they form and at which frequency. Recent reports indicate that DNA breaks stimulate DNA-RNA hybrid formation. We discuss the possibility that formation of hybrids may be the inevitable risk of breaks occurring at transcribed DNA regions. Such hybrids must be removed to allow repair, but it is unclear that they are intermediates required for repair.

DNA-RNA hybrids can be made between complementary RNA and DNA strands *in vitro* [1]. When this occurs in a double-stranded DNA (dsDNA), the RNA displaces the DNA strand of identical sequence and the whole structure formed by the DNA-RNA hybrid plus the displaced ssDNA is referred to as an R loop. *In vivo*, very-short DNA-RNA hybrids are formed inside the RNA polymerases (RNAPs) during transcription as well as with the RNA primers in Okazaki fragments, in both cases usually no more than 10 nucleotides long. Apart from that, longer DNA-RNA hybrids have been shown to form specifically during class switch recombination and during replication in mitochondrial DNA and in the bacterial plasmid ColE1, in all cases presumably as R loops (reviewed in [2]). The length of such hybrids can exceed 1 kb, as shown for the immunoglobulin (Ig) switch regions [3]. Their formation during such specific DNA metabolic processes, could in some cases be promoted by specialized proteins. Thus, the mechanism of action of the CRISPR-Cas9 system that relies on the identification of the target DNA site by the guide RNA that presumably form a DNA-RNA hybrid [4] is a good example of this. A different question is how DNA-RNA hybrids form

spontaneously throughout the genome likely without being mediated by a specialized protein system, becoming a threat to genome integrity and gene expression.

Preventing and dissolving R loops that compromise genome integrity

The spontaneous formation of R loops was first evidenced by the *in vivo* identification of DNA-RNA hybrids in yeast cells mutated in THO, a conserved complex that works at the interface of transcription and mRNP biogenesis [5] (**Fig. 1**). This led to a model in which specific RNA binding proteins (RBPs) protect the nascent RNA from hybridizing back with the DNA template. This model has gained support by the implication of an increasing number of RBPs in the prevention of R loop formation in yeast, *C. elegans* and human cells [6-11]. These R loops have typically been detected by their phenotypical consequences, such as genome instability, which is a stochastic phenomenon, and by *in vivo* detection methods that only provide measurements of relative increases rather than absolute values (reviewed in [2,12-14]). For this reason and despite well-established methods developed to analyze R-loops formed at the Ig switch regions [3,15], the length and frequency at which such spontaneous R loops are formed in the different mutant conditions from yeast to human cells is still unknown. This limits our capacity to clearly define the DNA features required for R loop formation during different DNA metabolic processes.

At present, we know a number of specific DNA features that are associated with a higher tendency to form R loops. Yeast and human cell lines show enrichment of DNA-RNA hybrids in the rDNA as well as in some repetitive regions such as yeast transposons, human L1 LINEs, telomeres and regions with chromatin modifications characteristic of repetitive DNA [16-21], which suggests that R loop formation is not necessarily random but can be stimulated by specific DNA characteristics. In this sense, although R-loops in the Ig switch regions only form co-transcriptionally [22], we cannot exclude that some RNAs might be able to re-anneal to the DNA *in trans,* once they have been released from the transcription site, despite the low stability of RNA due to its single strandedness [23]. This possibility might be particularly relevant in the case of small non-coding RNAs or RNAs generated from repetitive DNA sequences.

Cells possess different mechanisms to reduce the accumulation of harmful R loops. In addition to specific RBPs, such as the THO complex, that would act as a first option preventing the formation of DNA-RNA hybrids, specific enzymes such as RNase H1 and likely some RNA-DNA helicases, would act subsequently as a second option to remove the RNA once hybridized with the DNA (**Fig. 1**). Nevertheless, R loops can constitute a roadblock not only for the oncoming RNA polymerases but also for the RF,

thus being a major cause of transcription-replication conflicts that ultimately lead to genetic instability and chromosome fragility [13,24-31]. Consequently, a tight cellular coordination between transcription and replication is required so that upon replication fork (RF) stalling, cells do still have a third option to reduce R loops. The replication-related DNA repair machinery could contribute to R-loop dissolution (**Fig. 1**), as shown for the Fanconi Anemia and for BRCA1 and BRCA2 repair factors [30,32-34].

In addition, R-loops have been reported to induce antisense transcription, leading to the generation of dsRNA that in turn recruits the silencing machinery [35]. The antisense RNA could potentially counteract DNA-RNA hybrid formation by competing with the DNA strand in the hybrid. In support of this, an accumulation of DNA-RNA hybrids has been observed in the absence of DICER in *S. pombe*, although it has only been attributed to a defect in transcription termination [36]. Whether antisense RNA could represent a global mechanism to remove DNA-RNA hybrids is yet to be addressed. Anyhow, the existence diverse mechanisms to prevent, remove or dissolve DNA-RNA hybrids after the RF encounters an R loop, supports their potential harmfulness (**Fig. 1**). Indeed, R-loops might also be involved in neurological diseases and cancer (reviewed in [37]).

DNA breaks at transcribing DNA as promoters of DNA-RNA hybrid formation

Consistent with the high-energy cost that would support a single-stranded nucleic acid invading a closed dsDNA, several *in cis* elements have been shown to favor DNA-RNA hybridization (**Fig. 2**). The local negative supercoiling accumulated behind the elongating RNAP facilitates DNA strand opening, thus favoring invasion of the nascent RNA produced *in cis*, as can be inferred from genetic data in bacteria and yeast [26,38-40]. Also, DNA-RNA hybrids preferentially accumulate throughout the genome with a different likelihood depending on their GC or AT richness [18,19] and the formation of DNA secondary structures such as G quadruplexes formed at a nontemplate G-rich strand can also favour R loops as shown *in vitro* [22,41]. Indeed, given the complex and demanding step of ssDNA invasion into a dsDNA during homologous recombination (HR), which requires an available free DNA end and a number of proteins responsible for strand exchange (reviewed in [42]), it is unlikely that an internal sequence of a long ssRNA could stably hybridize with an intact dsDNA template spontaneously. Although a strand exchange protein does not seem to be required for R-loop formation, at least *in vitro*, a free DNA end would relieve any topological tension and enable greater breathing of the non-template DNA strand, so that the broken DNA would more easily undergo unwinding. In this scenario, the nascent RNA could better interact with the

template complementary DNA strand thus favouring the extension or stabilization of any putative transient RNA-DNA hybrid.

Importantly, it has been shown that an ssDNA nick in the Ig switch region transcribed from a T7 promoter strongly stimulates DNA-RNA hybrid formation [40]. Along this same line, camptothecin, which causes an ssDNA nick with its 3'-end covalently bound to Topo I (reviewed in [43]), also increases DNA-RNA hybrids [44,45]. An ssDNA nick in a transcribing double helix of DNA will release the topological constriction of the DNA, allowing DNA opening and rotation and the emergence of free ssDNA ends prone to anneal with the nascent RNA (**Fig. 2**). The relevance of DNA cleavage as one driving force in DNA-RNA hybrid formation thus seems clear.

Supporting the formation of DNA-RNA hybrids after DNA damage at transcription sites, the Calsou group detected a transcription-dependent accumulation of a catalytically inactive version of the *E. coli* RNase H to human DNA after laser micro-irradiation, which produces ssDNA nicks and DSBs [46]. Furthermore, they described the existence of a mechanism antagonizing R-loop formation after such damage that is dependent on all main players in the DNA damage response: ATM, ATR and DNA-PK. This mechanism relies on the exclusion of the RNA-bound pool of RNA-binding proteins from the damaged DNA [46]. Consistent with the general view that cleavage of the DNA behind an elongating RNAP facilitates DNA-RNA hybrids, two papers by the Godbout and Fisher groups have recently shown that DNA-RNA hybrids also accumulate upon double-strand break (DSB) induction [47,48].

The Godbout group reported previously that the human RNA-unwinding protein DEAD box 1 (DDX1) is an ATM target and forms foci at DSBs upon ionizing radiation (IR) [49]. Now, they have further shown that DDX1 depletion causes a delay in the repair of IR-induced γH2AX foci, reduced levels of RAD51 foci and increased lethality after IR, implying an involvement of DDX1 in the DNA damage response [47]. Importantly, IR-induced DDX1 foci are sensitive to RNase H treatment [49] as well as to the transcription inhibitors actinomycin D and cordycepin [47]. Using DNA-RNA Immuno-Precipitation (DRIP) as well as an *in vitro* assay based on the combined treatment with duplex-specific nuclease and/or RNase H, they have shown that DNA-RNA hybrids accumulate around a DSB when DDX1 is depleted [47]. Their results led them to propose that transcription promotes DDX1 recruitment to remove DNA-RNA hybrids at DSBs in human cells. The study constitutes new evidence that breakage of DNA during transcription facilitates the nascent RNA to hybridize back with the template DNA. In the same line, DDX1 foci are formed after RF breakage induced by long exposure to cisplatin or UV [49], which might reflect a high incidence of transcription-replication conflicts under such DNA damaging conditions.

On the other hand, the Fisher group has reported that in *S. pombe* RNase H is required for the repair of the DSBs induced by I-*Ppo*I endonuclease as well as for cell survival after cleavage [48]. This is accompanied by an increase in DNA-RNA hybrids at the site of DSB in *rnh1*Δ *rnh2*Δ cells. The results again strongly support that DNA breakage can lead to DNA-RNA hybrids. In this sense, it has been shown that the bacterial strand exchange protein RecA/Rad51 can catalyze RNA-DNA hybrid formation *in vitro* and its yeast ortholog Rad51 can modulate DNA-RNA hybrid formation *in trans* in *S. cerevisiae* cells [50-52]. RecA/Rad51 might sporadically contribute to make a cleaved DNA molecule more susceptible to hybrid formation, but this is unknown. In any case, the recent *S. pombe* study also shows that RNH1 overexpression affects DSB repair and survival without increasing DNA-RNA hybrids and leads to hyper-recombination between direct repeats when a DSB is induced between the repeats [48]. This suggests that the consequences of RNH1 action on RNA-DNA hybrids formed at broken DNA may not always be benign. Indeed, RNase H overexpression cannot compensate for the removal of RNA-DNA hybrids at DSBs in DDX1-depleted cells [47], implying that RNA removal does not always have straight and predictable consequences.

Possible impact of DNA-RNA hybrids on DNA repair

In light of these two reports [47,48] and the previous work of the Lieber and Calsou labs [40,46], we can conclude that ssDNA or dsDNA breakage, is a major driver of stable DNA-RNA hybrid formation and that DNA-RNA hybrids will have an impact on the repair of such breaks (**Fig. 2**). This adds a new step to the repair of breaks occurring in the DNA when this is being transcribed. In agreement with an important fraction of the genome prone to form DNA-RNA hybrids upon breakage, DDX1 depletion increases not only the sensitivity to IR but also to PARP inhibitors [47], a characteristic phenotype of cells deficient in HR factors. It is thus expected that the eventual formation of DNA-RNA hybrids at DSBs requires the removal of the RNA to allow DSB processing and repair. Consistently, a recent report suggests that the persistence of R-loops compromises DNA repair based on the fact that *rnh1*Δ *rnh2*Δ cells lead to a dramatic increase in unrepaired Rad52 foci and lethality in the absence of Top1 mostly due to defective break-induced replication at the repetitive rDNA region [53]. In principle, this should be specific to DNA regions undergoing transcription.

A breakage occurring at a transcribing DNA produces changes in local supercoiling and chromatin structure that, together with the free DNA ends generated, could facilitate hybridization of the nascent RNA with the ssDNA templates (**Fig. 2**).

The formation of DNA-RNA hybrids at the DSB site might also be facilitated by the 5'-end resection step of HR, because this generates a free 3'-ended DNA strand overhang that could easily hybridize with the nascent RNA (**Fig. 3**). In agreement with the non-resected template DNA strand forming a DNA-RNA hybrid, the amount of RPA-bound DNA after DSB induction is reduced in *rnh1*Δ *rnh2*Δ mutants and increased upon RNase H overexpression [48]. In this scenario, standard methods used to infer resection based on the exposure of ssDNA, including RPA, may not be valid.

Alternatively, DNA-RNA hybrids might be formed before resection, displacing the 5'-ended non-template strand, so that the major exposure of the 5'-ended ssDNA could promote its degradation by nucleases. Since 5'-end DSB resection is the key step channeling DSB repair towards HR [54], DNA-RNA hybrids at DSBs might therefore promote HR versus NHEJ repair. Indeed, the alteration by the DNA-RNA hybrid of the nature of the DSB end might impede its proper recognition by NHEJ factors (**Fig. 3**). Interestingly, and consistent with this view, a genome-wide analysis of repair of AsiSI nuclease-induced DSBs in human cells has shown that breaks at transcriptionally active chromatin are preferentially repaired by HR, whereas those transcriptionally inactive chromatin are repaired by NHEJ [55]. It would be interesting to know whether DNA-RNA hybrids contribute to this difference. As discussed above, for the repair to be productive, the RNA should be removed by RNase H1 or different DNA-RNA helicases. Interestingly, DDX1 seems to be required for DSB repair only after DNA resection [47].

Do DNA-RNA hybrids have a regular role in DNA repair?

The next question is whether DNA-RNA hybrids have a positive function in DSB repair. Transcription of non-coding RNA has been proposed to be important in the DDR upon DNA breakage in plants, human cells and flies [56-58]. Furthermore, RBPs and transcription factors have been shown in different studies to be enriched at damaged DNA although their direct role in DNA repair remains to be defined (reviewed in [59]). However, there is no experimental evidence to our knowledge for the requirement of DNA-RNA hybrids in the process of DSB repair. Indeed, the fact that DSBs can be repaired *in vitro* without any RNA suggests that R-loops are not obligatory. The detection of DNA-RNA hybrids at DSBs may just be a consequence of the nascent and sticky RNA being an unavoidable guest of breaks occurring at active chromatin. Examples of rare events that can happen during DSB repair include those cases in which an RNA molecule serve as an homologous template for recombination [60] without meaning that a template RNA is required in HR repair, or in which Ty and other mRNA-

derived cDNAs are used as intermediates in some DSB repair events [61,62], without meaning that such sequences enhance the repair process itself.

As indicated above, an important unresolved question is the absolute frequency at which R loops are formed at DSBs. A requirement for R loops in DSB repair would imply that they should be obligatory formed at DSBs, something for which so far we do not have evidence. In this sense, it is worth noting that the potential occurrence of DNA-RNA hybrids at breaks could be pretty general given the increasing evidence that supports that most of the genome in eukaryotes is transcribed at some moment, leading to mRNAs, rRNAs, tRNAs, snRNAs, or ncRNAs, including lncRNAs, scRNAs, CUTs, miRNAs or siRNAs. However, in an important fraction of the genome, transcription is spurious so that the frequency at which DSBs occur during transcription may vary greatly between chromosomal regions.

A certainly interesting observation is that RNAPII is enriched at DSB sites as observed by ChIP in *S. pombe* [48]. This does not necessarily mean that RNAPs are recruited to DSB sites. ChIPs only provide information about occupancy, not necessarily about active recruitment. Although speculative at this point, the DNA-RNA hybrid or the supercoiling and chromatin modifications associated with DSBs might block or retain a previously engaged RNAP at the vicinity of the breaks via different means, leading to an enhanced detection by ChIP of an RNAP present at the site prior to the break. In agreement, there are no changes in the transcript levels at the DSB site [48]. These and other intriguing questions would need to be solved in the near future. In this context, given that rDNA sequences are responsible for almost 50% of all DNA-RNA hybrids in yeast [19], an important question would be whether the enrichment at DSBs is restricted to RNAPII, or would it also be the case for RNAPI or RNAPIII polymerases.

In summary, RNA is a sticky molecule that can hybridize back with the DNA if circumstances are favorable. DNA breakage at transcribed regions leads to such favorable situations, leaving DNA defenseless with respect to the nascent RNA. Given that RNA is highly abundant and that most of the genome can be transcribed, the reports discussed here further support that understanding and identifying the factors and mechanisms that either prevent or eliminate DNA-RNA hybrids is necessary to have a complete understanding of genome dynamics, a subject of increasing relevance in cancer and genetic disease.

References

1. Thomas, M., White, R.L. & Davis, R.W. Hybridization of RNA to double-stranded DNA: formation of R-loops. *Proc. Natl. Acad. Sci. USA* **73**, 2294-8 (1976).
2. Aguilera, A. & Garcia-Muse, T. R loops: from transcription byproducts to threats to genome stability. *Mol. Cell* **46**, 115-24 (2012).
3. Yu, K., Chedin, F., Hsieh, C.L., Wilson, T.E. & Lieber, M.R. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.* **4**, 442-51 (2003).
4. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-21 (2012).
5. Huertas, P. & Aguilera, A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* **12**, 711-21 (2003).
6. Li, X. & Manley, J.L. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* **122**, 365-78 (2005).
7. Paulsen, R.D. et al. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol. Cell* **35**, 228-39 (2009).
8. Castellano-Pozo, M., Garcia-Muse, T. & Aguilera, A. R-loops cause replication impairment and genome instability during meiosis. *EMBO Rep.* **13**, 923-9 (2012).
9. Wahba, L., Amon, J.D., Koshland, D. & Vuica-Ross, M. RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol. Cell* **44**, 978-88 (2011).
10. Stirling, P.C. et al. R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes Dev* **26**, 163-75 (2012).
11. Sridhara, S. et al. Transcription dynamics prevents RNA-mediated genomic instability through SRPK2-dependent DDX23 phosphorylation. *Cell Rep.* **18**, 334-343 (2017).
12. Skourti-Stathaki, K. & Proudfoot, N.J. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev.* **28**, 1384-96 (2014).
13. Sollier, J. & Cimprich, K.A. Breaking bad: R-loops and genome integrity. *Trends Cell. Biol.* **25**, 514-22 (2015).
14. Santos-Pereira, J.M. & Aguilera, A. R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.* **16**, 583-97 (2015).
15. Huang, F.T. et al. Sequence dependence of chromosomal R-loops at the immunoglobulin heavy-chain Smu class switch region. *Mol. Cell. Biol.* **27**, 5921-32 (2007).
16. Chan, Y.A. et al. Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. *PLoS Genet.* **10**, e1004288 (2014).
17. El Hage, A., Webb, S., Kerr, A. & Tollervey, D. Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet.* **10**, e1004716 (2014).
18. Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I. & Chedin, F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell* **45**, 814-25 (2012).

19. Wahba, L., Costantino, L., Tan, F.J., Zimmer, A. & Koshland, D. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev.* **30**, 1327-38 (2016).
20. Nadel, J. et al. RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. *Epigenetics Chromatin* **8**, 46 (2015).
21. Balk, B., Dees, M., Bender, K. & Luke, B. The differential processing of telomeres in response to increased telomeric transcription and RNA-DNA hybrid accumulation. *RNA Biol.* **11**, 95-100 (2014).
22. Duquette, M.L., Handa, P., Vincent, J.A., Taylor, A.F. & Maizels, N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.* **18**, 1618-29 (2004).
23. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709-15 (1993).
24. Wellinger, R.E., Prado, F. & Aguilera, A. Replication fork progression is impaired by transcription in hyperrecombinant yeast cells lacking a functional THO complex. *Mol. Cell. Biol.* **26**, 3327-34 (2006).
25. Gan, W. et al. R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.* **25**, 2041-56 (2011).
26. Tuduri, S. et al. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat. Cell Biol.* **11**, 1315-24 (2009).
27. Helmrich, A., Ballarino, M. & Tora, L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol. Cell* **44**, 966-77 (2011).
28. Houlard, M., Artus, J., Leguillier, T., Vandormael-Pournin, S. & Cohen-Tannoudji, M. DNA-RNA hybrids contribute to the replication dependent genomic instability induced by Omcg1 deficiency. *Cell Cycle* **10**, 108-17 (2011).
29. Merrikh, H., Zhang, Y., Grossman, A.D. & Wang, J.D. Replication-transcription conflicts in bacteria. *Nat. Rev. Microbiol.* **10**, 449-58 (2012).
30. Schwab, R.A. et al. The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Mol. Cell* **60**, 351-61 (2015).
31. Garcia-Muse, T. & Aguilera, A. Transcription-replication conflicts: how they occur and how they are resolved. *Nat. Rev. Mol. Cell. Biol.* (2016).
32. Bhatia, V. et al. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* **511**, 362-5 (2014).
33. Hatchi, E. et al. BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol. Cell* **57**, 636-47 (2015).
34. Garcia-Rubio, M.L. et al. The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. *PLoS Genet* **11**, e1005674 (2015).
35. Skourti-Stathaki, K., Kamieniarz-Gdula, K. & Proudfoot, N.J. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature* **516**, 436-9 (2014).
36. Castel, S.E. et al. Dicer promotes transcription termination at sites of replication stress to maintain genome stability. *Cell* **159**, 572-83 (2014).

37. Groh, M. & Gromak, N. Out of balance: R-loops in human disease. *PLoS Genet.* **10**, e1004630 (2014).
38. Drolet, M. Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology. *Mol. Microbiol.* **59**, 723-30 (2006).
39. El Hage, A., French, S.L., Beyer, A.L. & Tollervey, D. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev.* **24**, 1546-58 (2010).
40. Roy, D., Zhang, Z., Lu, Z., Hsieh, C.L. & Lieber, M.R. Competition between the RNA transcript and the nontemplate DNA strand during R-loop formation in vitro: a nick can serve as a strong R-loop initiation site. *Mol. Cell. Biol.* 30, 146-59 (2010).
This study provides the first evidence that a nick in a dsDNA induces DNA-RNA hybrid formation in an *in vitro* transcribed supercoiled plasmid containing the Ig S repeats.
41. Duquette, M.L., Huber, M.D. & Maizels, N. G-rich proto-oncogenes are targeted for genomic instability in B-cell lymphomas. *Cancer Res.* **67**, 2586-94 (2007).
42. Heyer, W.D., Ehmsen, K.T. & Liu, J. Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.* **44**, 113-39 (2010).
43. Pourquier, P. & Pommier, Y. Topoisomerase I-mediated DNA damage. *Adv Cancer Res* **80**, 189-216 (2001).
44. Sordet, O. et al. Ataxia telangiectasia mutated activation by transcription- and topoisomerase I-induced DNA double-strand breaks. *EMBO Rep.* **10**, 887-93 (2009).
45. Groh, M., Lufino, M.M., Wade-Martins, R. & Gromak, N. R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. *PLoS Genet.* **10**, e1004318 (2014).
46. Britton, S. et al. DNA damage triggers SAF-A and RNA biogenesis factors exclusion from chromatin coupled to R-loops removal. *Nucleic Acids Res.* 42, 9047-62 (2014).
This study shows that a catalytically inactive version of the *E. coli* RNase H is accumulated in a transcription-dependent manner at human DNA after laser micro-irradiation, which produces ssDNA nicks and DSBs.
47. Li, L. et al. DEAD Box 1 Facilitates Removal of RNA and Homologous Recombination at DNA Double-Strand Breaks. *Mol. Cell. Biol.* 36, 2794-2810 (2016).
This study reports DNA-RNA hybrid accumulation around a DSB when the helicase DDX1 is depleted in human cells.
48. Ohle, C. et al. Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair. *Cell* 167, 1001-1013 e7 (2016).
This study reports that DNA-RNA hybrids accumulate at *I-Ppo*I-induced DSBs in *rnh1*Δ *rnh2*Δ *S. pombe* cells and that RNAPII is enriched at the DSB site. The authors propose that DNA-RNA hybrids are required for DSB repair.
49. Li, L., Monckton, E.A. & Godbout, R. A role for DEAD box 1 at DNA double-strand breaks. *Mol. Cell. Biol.* **28**, 6413-25 (2008).
50. Kasahara, M., Clikeman, J.A., Bates, D.B. & Kogoma, T. RecA protein-dependent R-loop formation in vitro. *Genes Dev.* **14**, 360-5 (2000).

51. Zaitsev, E.N. & Kowalczykowski, S.C. A novel pairing process promoted by Escherichia coli RecA protein: inverse DNA and RNA strand exchange. *Genes Dev.* **14**, 740-9 (2000).
52. Wahba, L., Gore, S.K. & Koshland, D. The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. *Elife* **2**, e00505 (2013).
53. Amon, J.D. & Koshland, D. RNase H enables efficient repair of R-loop induced DNA damage. *Elife* 5(2016).
This study suggests that persistent R-loops compromise DNA repair in *S. cerevisiae* cells. It shows a dramatic increase in unrepaired Rad52 foci in *rnh1*Δ *rnh2*Δ cells and that such cells are inviable in the absence of Top1, this inviability being mainly caused by defective break-induced replication at the rDNA.
54. Huertas, P. DNA resection in eukaryotes: deciding how to fix the break. *Nat. Struct. Mol. Biol.* **17**, 11-6 (2010).
55. Aymard, F. et al. Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat. Struct. Mol. Biol.* 21, 366-74 (2014).
This study, based on a genome-wide analysis in human cells, suggests that restrictase-induced DSBs at transcribed euchromatin are preferentially repaired via HR, wheras those in non-transcribed euchromatin, preferentially via NHEJ.
56. Wei, W. et al. A role for small RNAs in DNA double-strand break repair. *Cell* **149**, 101-12 (2012).
57. Francia, S. et al. Site-specific DICER and DROSHA RNA products control the DNA-damage response. *Nature* **488**, 231-5 (2012).
58. Michalik, K.M., Bottcher, R. & Forstemann, K. A small RNA response at DNA ends in Drosophila. *Nucleic Acids Res.* **40**, 9596-603 (2012).
59. D'Alessandro, G. & di Fagagna, F.D. Transcription and DNA damage: holding hands or crossing swords? *J. Mol. Biol.* (2016).
60. Keskin, H. et al. Transcript-RNA-templated DNA recombination and repair. *Nature* **515**, 436-9 (2014).
61. Moore, J.K. & Haber, J.E. Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. *Nature* **383**, 644-6 (1996).
62. Yang, J., Zimmerly, S., Perlman, P.S. & Lambowitz, A.M. Efficient integration of an intron RNA into double-stranded DNA by reverse splicing. *Nature* **381**, 332-5 (1996).

Acknowledgements

We thank Diane Haun for style correction. Research in AA's lab is funded by the European Research Council, Spanish Ministry of Economy and Competitiveness, Junta de Andalucía and World Cancer Research. B.G-G is a postdoctoral fellow of the Scientific Foundation of the Spanish Association Against Cancer (AECC).

FIGURE LEGENDS

Figure 1. Multiple protein factors act at three different steps to prevent harmful DNA-RNA hybrid accumulation

RNA-binding proteins (grey circles) directly associate with the nascent RNA (red) to prevent the formation of DNA-RNA hybrids during transcription by RNA polymerase II (RNAPII), as occurs during mRNP biogenesis (middle panel). If formed, DNA-RNA hybrids may be removed by RNAse H or by a subset of RNA-DNA helicases before genome integrity is compromised. However, if DNA-RNA hybrids persist, they can also be dissolved during replication via replication-associated DNA repair factors such as BRCA1, BRCA2 and the Fanconi anemia (FA) pathway.

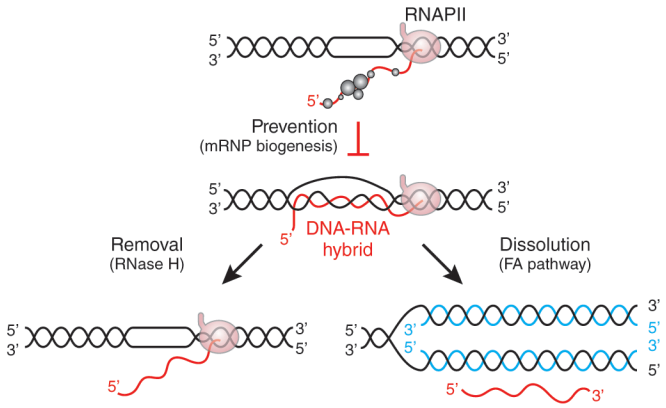
Figure 2. Structural features/elements that stimulate DNA-RNA hybrid formation along transcribed DNA

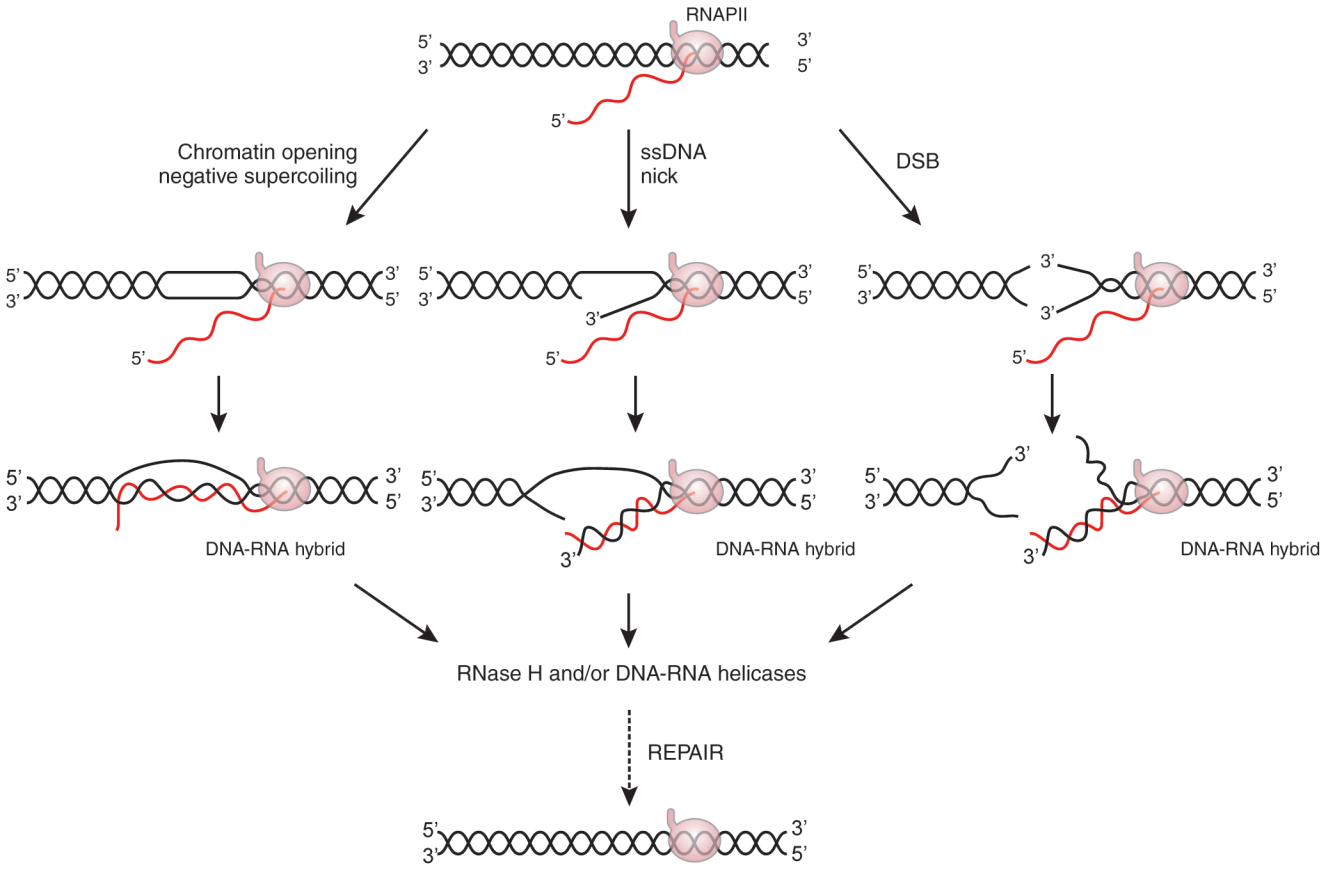
Different nucleic acid and chromatin structures favor DNA hybridization of the nascent RNA within transcribed chromatin. For example, chromatin opening, the accumulation of negative supercoiling or suboptimal mRNP formation can stimulate co-transcriptional DNA-RNA hybridization (left panel). In addition, DNA breaks, either single-stranded (middle panel) or double-stranded (right panel), can provide a free 3'-DNA end that may favor DNA hybridization with the nascent RNA (red). In such cases, the resulting DNA-RNA hybrid would need to be removed by the action of RNAseH or DNA-RNA helicases (bottom) to leave the ssDNA free and allow the additional steps (dashed arrow) required for repair of the DNA break.

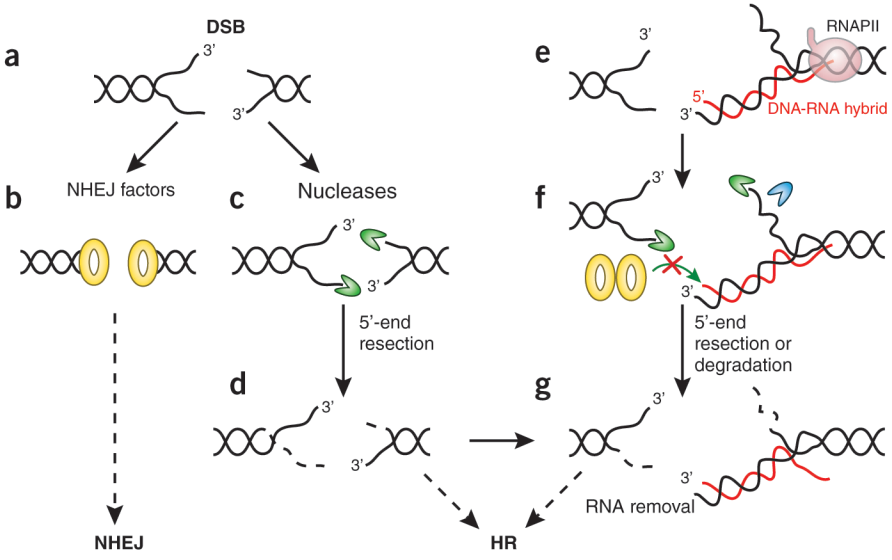
Figure 3. Potential impact of DNA-RNA hybrids on DNA DSB repair

When a DSB occurs (a), the NHEJ factors Ku70/80 heterodimer (yellow ring) directly bind and protect the DSB ends from resection, and channel repair towards NHEJ (b). Alternatively, when the 5' ends of the break are resected by specific nucleases (green) (c,d), DSBs are repaired via HR. Within transcribed

regions, the 3’-end DNA strand of the DSB is the most likely to hybridize with the nascent RNA. The most likely scenario, as depicted, show a nascent RNA coming from a transcribing RNA polymerase II (RNAPII) travelling outward from the site of the DSB. The DSB upstream of the transcribing RNAPII facilitates nascent RNA hybridization with the 3’-ended template DNA strand (e). The displaced 5’-end DNA strand would be less protected from degradation by nucleases (green and blue), including those normally involved in resection (f). In addition, the alteration of the DSB end (from DNA-DNA to DNA-RNA plus ssDNA) could make it less recognizable by NHEJ factors (f). It is therefore possible that a DNA-RNA hybrid at a DSB could favor 5’-end resection (g), channeling repair preferentially to HR. Eventually, a DNA-RNA hybrid could also be formed in a transcribed region that undergoes a DSB, after the DNA resection step (from d to g) with similar consequences. Once the DSB is resected at its 5’-end, it is committed to repair by HR via subsequent steps (dashed arrow). In the case of DNA-RNA hybrid has been formed, it would need to be removed first to allow HR.


RNAPII
5'
3'
3'
5'
5'
Prevention
(mRNP biogenesis)
5'
3'
3'
5'
5'
DNA-RNA
hybrid
Removal
(RNase H)
Dissolution
(FA pathway)
5'
3'
3'
5'
5'
5'
3'
5'
3'
3'
5'
3'
5'
5'
3'


RNAPII
5'
3'
3'
5'
5'
Chromatin opening
negative supercoiling
ssDNA
nick
DSB
3'
3'
5'
DNA-RNA hybrid
DNA-RNA hybrid
DNA-RNA hybrid
RNase H and/or DNA-RNA helicases
REPAIR


DSB
a
3'
3'
NHEJ factors
b
Nucleases
c
3'
3'
5'-end
resection
d
3'
3'
NHEJ
HR
e
3'
RNAPII
5'
3'
DNA-RNA hybrid
f
3'
3'
5'-end
resection or
degradation
g
3'
3'
RNA removal