Published in final edited form as:

DNA Repair (Amst). 2013 March 1; 12(3): 205–211. doi:10.1016/j.dnarep.2012.12.004.

Two distinct mechanisms of Topoisomerase 1-dependent mutagenesis in yeast

Jang-Eun Cho^a, Nayun Kim^a, Yue C. Li^a, and Sue Jinks-Robertson^{a,1}

Jang-Eun Cho: jangeun.cho@duke.edu; Nayun Kim: nayun.kim@duke.edu; Yue C. Li: claire.li@duke.edu; Sue Jinks-Robertson: sue.robertson@duke.edu

^aDepartment of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA

Abstract

Topoisomerase 1 (Top1) resolves transcription-associated supercoils by generating transient single-strand breaks in DNA. Top1 activity in yeast is a major source of transcription-associated mutagenesis, generating a distinctive mutation signature characterized by deletions in short, tandem repeats. A similar signature is associated with the persistence of ribonucleoside monophosphates (rNMPs) in DNA, and it also depends on Top1 activity. There is only partial overlap, however, between Top1-dependent deletion hotspots identified in highly transcribed DNA and those associated with rNMPs, suggesting the existence of both rNMP-dependent and rNMP-independent events. Here, we present genetic studies confirming that there are two distinct types of hotspots. Data suggest a novel model in which rNMP-dependent hotspots are generated by sequential Top1 reactions and are consistent with rNMP-independent hotspots reflecting processing of a trapped Top1 cleavage complex.

Keywords

Transcription-associated mutagenesis; topoisomerase 1; mutagenesis; ribonucleotide

1. Introduction

Spontaneous mutagenesis is a contributor to the development of somatic diseases as well as an underlying mechanism for evolutionary processes [1]. The rate of mutagenesis is modulated by transcription, with analysis in *Saccharomyces cerevisiae* demonstrating a proportional relationship between transcription and mutagenesis [2]. Studies of transcription-associated mutagenesis (TAM) in yeast have revealed several contributors to this phenomenon. First, elevated DNA damage associated with active transcription can stimulate mutagenesis [3, 4]. Such damage may be facilitated by the formation of R-loops in which the transcript remains stably base paired with the DNA template (reviewed in [5]).

© 2012 Elsevier B.V. All rights reserved.

¹Corresponding Author: Sue Jinks-Robertson, Department of Molecular Genetics and Microbiology, Duke University Medical Center, 384 CARL Bldg. DUMC 3020, 213 Research Dr. Durham, NC 27710, Telephone: +1-919-681-7273, Fax: +1-919-681-7273, sue.robertson@duke.edu.

Conflict of interest

The authors declare that there is no conflict of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

This exposes the non-transcribed strand as a single-strand DNA region, which is expected to be more sensitive to endogenous DNA damaging agents than duplex DNA [6]. Second, increased dUTP is incorporated during active transcription and its removal can initiate TAM [7]. Finally, the major contributor to TAM in an unbiased forward mutation assay is activity of Topoisomerase I (Top1) [8, 9]. Top1 produces a very distinctive mutation signature, 2–5 base pair (bp) deletions at short, tandem repeats, in which deletion size coincides with the repeat-unit size.

Transcription is associated with an accumulation of positive supercoils ahead of and negative supercoils behind the RNA polymerase complex [10]. In *S. cerevisiae*, such transcription-associated supercoils are primarily resolved by Top1 [11], which is recruited by interaction with the phosphorylated carboxy-terminal domain of RNA polymerase II [12]. Top1 binds to duplex DNA and cleaves one of the strands, forming a covalent phosphotyrosine bond with the 3'-end of the nicked DNA and generating a 5'-hydroxyl group (5'-OH) on the other side of the nick. Rotation of Top1 around the intact strand relieves the torsional stress, and Top1-mediated re-ligation using the 5'-OH restores the integrity of the cleaved strand. The covalent Top1-DNA intermediate is referred to as a Top1 cleavage complex (Top1cc) [13]. We previously suggested that irreversible trapping of the Top1cc can initiate the formation of the signature short deletions at tandem repeats [8]. In this model, processing of the Top1cc generates a small gap, which is converted to a ligatable nick by repeat-mediated realignment of the complementary strands (Fig. 1).

Top1-dependent deletions are also stimulated by ribonucleoside monophosphates (rNMPs) in DNA [14]. Replicative polymerases, such as Pol ϵ , can incorporate rNMPs into DNA [15], and RNase H2 initiates removal of such rNMPs in yeast [16–18]. In an RNase H2-defective background, a mutation signature identical to that of Top1-dependent TAM has been observed [19]. Significantly, these rNMP-associated deletions also result from activity of Top1 [14]. Such deletions, however, are unlikely to reflect processing of a trapped Top1cc. Following incision at an rNMP, the enzyme can be released by attack of the 3'-phosphotyrosyl bond by the 2'-OH of ribose, leaving a 2'-3' cyclic phosphate on one side of the nick and a 5'-OH on the other [14, 20]. We previously suggested that, as with a trapped Top1cc, rNMP-dependent deletions reflect processing of the non-canonical ends created by this reaction (Fig. 1).

In the current study, we investigate the combined effect of active transcription and rNMPs in DNA on Top1-dependent deletions. Through use of a mutant Top1 that retards the re-ligation reaction and a mutant Pol ϵ that inserts less rNMPs into DNA than WT, we provide further evidence that Top1-dependent deletions reflect the repair of two distinct types of Top1-generated DNA lesions: a trapped Top1cc and a nick generated by Top1 incision at an rNMP. A new model for the Top1-dependent, rNMP-associated deletion process is proposed that requires two sequential Top1 cleavage reactions and depends on the re-ligation activity of Top1.

2. Material and Methods

2.1 Strain construction

Yeast strains for the chromosomal *CAN1* forward and *lys2* reversion assays were descendents of SJR282 (*MATa ade2-101oc his3 Δ 200 ura3 Δ Nco suc2 gal80 Δ ::HIS3*; [3]) and YPH45 (*MATa ura3-52 ade2-101oc trp1 Δ 1*; [21]), respectively. The WT *pGAL-CAN1*, pLYS-lys2 Δ A746,NR,(AT)₂, pTET-lys2 Δ A746,NR,(AT)₂, pLYS-lys2 Δ A746,NR,(TC)₃, pTET-lys2 Δ A746,NR,(TC)₃, pLYS-lys2 Δ A746,NR,(AG)₄, pTET-lys2 Δ A746,NR,(AG)₄, pLYS-lys2 Δ A746, pTET-lys2 Δ A746 and pTET-lys2 Δ Bgl strains were previously described [2, 8, 14]. RNH201 and/or TOP1 were deleted by one-step allele replacement

using PCR-generated deletion cassettes amplified from plasmids containing appropriate selectable markers. *POL2* was replaced with *pol2-M644L* using a two-step allele replacement strategy [17]. A complete list of strains is given in Table S1.

2.2 Mutation rates and spectra

To determine mutation rates, cells were nonselectively grown to saturation in YEP medium (1% yeast extract, 2% Bacto-peptone, and 250 $\mu\text{g/mL}$ adenine hemisulfate with 2% agar for plates) supplemented with 2% glycerol and 2% ethanol (YEPGE); 2 $\mu\text{g/mL}$ doxycycline (Sigma) was added as appropriate. Dilutions of individual cultures were plated on YEPD to determine total cell numbers or on synthetic complete dextrose medium lacking lysine (SCD-Lys) to determine the numbers of Lys⁺ revertants. Canavanine-resistant colonies were selected on SCD-Arg plates containing 60 $\mu\text{g/mL}$ L-canavanine sulfate. Mutation rates were calculated using the method of the median [22] and 95% confidence intervals were determined as described previously [23].

To obtain mutation spectra, the region of interest of was PCR-amplified, and the resulting fragments were sequenced by the Duke University DNA Analysis Facility. The rate of a specific type of mutation was calculated by multiplying its proportion in the spectrum by the total Lys⁺ rate.

2.3 Mutation frequencies with Top1-T722A expression

CEN plasmids expressing Top1-T722A or WT Top1 contained *URA3* as a selectable marker, and both proteins were expressed from the *CUP1* promoter. Transformants were selected on SCD-Ura plates, and independent cultures were directly started from individual colonies without prior purification. Cells were grown in SCD-Ura medium for three days and then plated on SCD-Ura to determine plasmid retention and on SCD-Ura-Lys to obtain the Lys⁺ reversion frequency. Addition of copper to induce transcription from *CUP1* was unnecessary in our strain background. At least six independent cultures were started for each strain, and each primary transformation was done twice. Because the starting cell numbers were not same and cultures were at different cell densities when plated, median reversion frequencies rather than rates were determined. Revertants were sequenced as above to determine the 2-bp deletion frequency and 95% confidence intervals of the median frequency of 2-bp deletions were calculated [24].

3. Results

3.1 Combined effect of active transcription and RNase H2 loss on Top1-dependent deletions at *CAN1*

We previously reported that Top1-dependent, 2–5 bp deletions at the *CAN1* locus arise under two distinct conditions: (1) when transcription is highly activated from a galactose-regulated promoter in a WT background and (2) when rNMPs fail to be removed from DNA in an RNase H2-defective (*rnh201Δ*) background [8, 14]. To compare how RNase H2 loss affects the rate of deletions under low-versus high-transcription conditions, we deleted *RNH201* in a strain containing *CAN1* either regulated by its own promoter or by the highly activated *GAL* promoter (*pGAL*). The rate of 2–5 bp deletions increased about 1600-fold under high- relative to low-transcription conditions when RNase H2 was active (Fig. 2A). Similar to the 10-fold increase observed under low-transcription conditions, loss of RNase H2 elevated the total rate of deletions an additional 13-fold under high-transcription conditions, and all deletions were Top1 dependent. The observed multiplicative effect of transcription and RNase H2 loss is expected if the level of persistent rNMPs in *CAN1* is the same under high- and low-transcription conditions in an *rnh201Δ* background, and if the

stimulatory effect of transcription reflects recruitment of Top1 to remove associated supercoils.

3.2 Effect of active transcription and RNase H2 loss on isolated *CAN1* hotspots

Similar to the pattern observed when the low-transcription *rnh201Δ* and high-transcription *RNH201* spectra were compared [14], the positions of 2-bp deletion hotspots at *CAN1* only partially overlapped when the high-transcription spectra generated from the *RNH201* and the *rnh201Δ* backgrounds were compared (Fig. 2B). The variable effects of RNase H2 loss on Top1-dependent hotspots under high-transcription conditions support our previous interpretation that there are likely two classes of hotspots: those reflecting irreversible Top1 cleavage at an rNMP and those reflecting an rNMP-independent Top1 cleavage intermediate.

In previous studies, three of the Top1-dependent 2-bp deletion hotspots [(AT)₂, (TC)₃, and (AG)₄] identified at *CAN1* were individually transplanted as ~30 bp fragments into the “reversion window” of a chromosomal *lys2ΔA746,NR* frameshift allele [8, 14]. The *lys2ΔA746,NR* allele contains a 1-bp deletion, and reversion to lysine prototrophy occurs via a compensatory net 1-bp insertion [25]. Importantly, this allele has a very low background reversion rate and thus can efficiently detect 2-bp deletions at Top1-dependent hotspots, allowing each to be studied in isolation. To more accurately assess the combined effects of RNase H2 loss and transcription at individual hotspots, we deleted *RNH201* in strains containing the low (*pLYS*)- or high (*pTET*)-transcription *lys2ΔA746,NR*,(AT)₂, *lys2ΔA746,NR*,(TC)₃, or *lys2ΔA746,NR*,(AG)₄ allele. *pTET* is a tetracycline/doxycycline repressible promoter and drives high levels of transcription when cells are grown in the absence of doxycycline [2].

In the presence of RNase H2, (AG)₄ was not detected as a transcription-associated deletion hotspot when the *pGAL-CAN1* allele was examined (Fig. 2B). When analyzed in isolation, however, a 28-fold stimulatory effect of transcription on 2-bp deletions within the (AG)₄ tandem repeat was evident (Fig. 3A). Due to the absence of events under low-transcription conditions, the precise stimulatory effect of transcription on the (TC)₃ and (AT)₂ hotspots within *CAN1* could not be determined [8]. This was possible with the transplanted hotspots, however, and we observed a 95- and 71-fold transcription-associated increase, respectively, in the 2-bp deletion rates (Fig. 3 B–C). Consistent with observations made under low-transcription conditions [14], deletion of *RNH201* had very different effects on deletions within the individual (AT)₂, (TC)₃ and (AG)₄ hotspots under high-transcription conditions. While RNase H2 loss had no significant effect on the rate of deletions at the (AT)₂ hotspot, those at the (TC)₃ hotspot increased an additional 7.6-fold and those at the (AG)₄ hotspot increased 870-fold.

In the context of *pGAL-CAN1*, loss of RNase H2 was associated with a similar increase (approximately 10-fold; see section 3.1) in the overall rates of 2–5 bp deletions under either low- or high-transcription conditions. A comparable effect of *RNH201* deletion under different transcription conditions did not, however, necessarily extend to the isolated hotspots. In particular, RNase H2 loss had a larger effect on deletions at the (AG)₄ hotspot under high- than under low-transcription conditions (870- and 240-fold, respectively). One interpretation of this result is that transcription has an effect on rNMP-dependent hotspots that extends beyond simply recruiting more Top1.

3.3 Reduced incorporation of rNMPs into DNA decreases deletions at rNMP-associated hotspots

The behavior of isolated deletion hotspots upon *RNH201* loss under low- or high-transcription conditions led us to speculate that the (TC)₃ and (AG)₄ hotspots are rNMP-associated deletion hotspots, while (AT)₂ is an rNMP-independent hotspot. An alternative explanation, however, is that the (AT)₂ hotspot is also rNMP-dependent, and the relevant rNMP is simply refractory to RNase H2 activity. To provide more direct evidence for two distinct types of hotspots, we replaced WT Pol2, the catalytic subunit of Pol ϵ , with a mutant form (Pol2-M644L) that incorporates fewer rNMPs into genomic DNA than the WT enzyme [17]. The rationale behind this approach is that if deletions at a given hotspot are initiated by Top1 cleavage at an rNMP, then the frequency of these events should decrease when fewer rNMPs are incorporated into DNA. If rNMPs are irrelevant to hotspot activity, however, then the deletion rate should not be affected by the replacement of Pol2 with Pol2-M644L. Expression of Pol2-M644L in an *rnH201* Δ background decreased the rate of 2-bp deletions at the (AG)₄ hotspot about 40-fold, and those at (TC)₃ hotspot about 4-fold (Fig. 4A), indicating that most deletions at these hotspots are initiated by Top1 cleavage at an rNMP normally inserted by Pol ϵ . By contrast, the rates of deletions at (AT)₂ were comparable in strains expressing WT Pol2 or Pol2-M644L, supporting the interpretation that deletions at (AT)₂ arise in a predominantly, if not exclusively, rNMP-independent manner.

3.4 Reduced re-ligation activity of Top1 has different effects on rNMP-associated and rNMP-independent hotspots

We previously suggested that rNMP-independent deletions reflect repair of a stably trapped Top1cc [14]. The mutant Top1-T722A protein has been shown to have reduced re-ligation activity *in vitro* [26] [27], which is expected to increase the half-life of the Top1cc *in vivo* and thereby promote more trapping. We predicted that if rNMP-independent deletions at the (AT)₂ hotspot were produced through processing of a trapped Top1cc, then expression of Top1-T722A should increase the frequency of these deletions. We thus analyzed the effect of Top1-T722A expression on the frequency of deletions at the transplanted (AT)₂ hotspot by transforming an *rnH201* Δ *top1* Δ strain with an empty plasmid, a plasmid expressing WT Top1 or a plasmid expressing Top1-T722A. As predicted, the frequency of 2-bp deletions at the (AT)₂ hotspot was 11-fold higher with Top1-T722A expression than with WT Top1 expression (Fig. 4B).

In vitro data have demonstrated that vaccinia Top1 cleaves at an rNMP to form a transient Top1cc, and then is released from the DNA when the phosphotyrosine bond of the Top1cc is attacked by the vicinal 2'-OH group [20]. We proposed that, as with a trapped Top1cc, processing the resulting cyclic phosphate end would lead to formation of a small gap within the tandem repeat [14]. Because rNMP-dependent hotspots do not involve formation of a trapped Top1cc, we predicted that deletions at the (TC)₃ and (AG)₄ hotspots would occur at a similar level in *rnH201* Δ *top1* Δ cells expressing either Top1-T722A or WT Top1. As shown in Fig. 4B, however, the frequency of 2-bp deletions at the (TC)₃ and (AG)₄ hotspots were 6-fold and 8-fold lower, respectively, with Top1-T722A expression than with WT Top1 expression. The opposing effect of Top1-T722A expression on deletions at isolated hotspots confirms that there are two distinctive classes of Top1-dependent deletion hotspots and furthermore suggests that the re-ligation activity of Top1 is important for deletion formation at rNMP-associated hotspots.

3.5 Effect of active transcription and increased rNMPs on a 4-bp deletion hotspot

We previously described a Top1-dependent, rNMP-associated 4-bp deletion hotspot [(AGCT)₂] in the *lys2* Δ *Bgl* reversion assay under low-transcription conditions [14]. The chromosomal *lys2* Δ *Bgl* allele contains a +1 frameshift mutation and reverts to lysine

prototrophy via a compensatory, net 1-bp deletion within a reversion window that coincides with that of the *lys2ΔA746,NR* allele [28]. To examine the combined effect of high transcription and RNase H2 loss on deletions at the (AGCT)₂ hotspot, reversion of a *pTET-lys2ΔBgl* allele was examined in the presence or absence of doxycycline (low- and high-transcription conditions, respectively). Loss of RNase H2 was associated with a 26-fold increase in hotspot activity under low-transcription conditions, and 110-fold increase under high-transcription conditions (Fig. 5A). Although there also may be a greater-than-multiplicative effect of active transcription and RNase H2 loss on the 4-bp deletions, the effect of transcription alone could not be accurately gauged because only one 4-bp deletion was observed in the spectrum under high-transcription conditions in the WT background.

3.6 rNMP-dependent deletions at a 6A hotspot are at least 2 bp

In an earlier TAM study, we described an increase in 2-bp deletions at a 6A mononucleotide run within the chromosomal *lys2ΔA746* reversion window under high-transcription conditions (*pTET-lys2ΔA746*, [2]). This allele is the parent of the *lys2ΔA746,NR* (NR, no-run) allele, which was constructed by eliminating all mononucleotide runs longer than 3N from the *lys2ΔA746* reversion window [29]. When initially described, the mechanism for generating the transcription-associated, 2-bp deletions at the 6A run was not known. Because the 6A run is a tandem repeat, which seems to be a necessary condition for Top1-dependent deletions, we hypothesized that 2-bp deletions at this position might similarly require Top1. The deletion of *TOP1* completely eliminated 2-bp deletions at the 6A run, indicating that the 6A run is indeed an additional Top1-dependent deletion hotspot (Fig. 5B). Furthermore, deletion of *RNH201* strongly increased the rate of deletions at the 6A run under either low- or high-transcription conditions, demonstrating that the 6A run is also an rNMP-associated deletion hotspot.

The 6A run in the *pTET-lys2ΔA746* reversion window can be considered an (AA)₃ dinucleotide repeat. When viewed in this manner, it is striking that all Top1-dependent deletion hotspots we have observed to date are comprised of repeat units that are at least 2 bp. The 6A hotspot provides a unique opportunity to examine whether Top1-generated deletions of only a single base pair also are possible. We thus monitored deletions at this position using the *pTET-lys2ΔBgl* allele, which detects net 1-bp deletions (Fig. S1). One-bp deletions in the 6A run are, in fact, the most common event that reverts the *lys2ΔBgl* allele in *RNH201* backgrounds, comprising 34% and 18% of the reversion spectrum under high- and low-transcription conditions, respectively [14]. Although high transcription enhanced 1-bp deletions in the 6A run by 47-fold, neither *RNH201* deletion nor *TOP1* deletion had a significant effect on these events (Fig. 5A). This indicates that Top1 cleavage at rNMPs only rarely generates 1-bp deletions, if it does so at all, and suggests that the minimum size of efficient Top1-dependent, rNMP-associated deletions is likely 2 bp.

4. Discussion

The separate examination of the effect of transcription and RNase H2 loss on Top1-dependent deletions previously led us to propose that there are two distinct classes of these events: rNMP-associated deletions initiated by Top1 incision at an rNMP and rNMP-independent deletions initiated by Top1 trapping. Here, we tested the combined effect of the high transcription and RNase H2 loss on Top1-dependent deletions at *CAN1*, on three individual 2-bp deletion hotspots originally identified at *CAN1* and on the endogenous 4-bp and 2-bp deletion hotspots found within the *lys2* reversion window. We additionally examined the effect of rNMP-restrictive Pol2-M644L or ligation-deficient Top1-T722A on deletions at representative hotspots. Below we discuss the results obtained in these analyses and propose a new model for the Top1-dependent, rNMP-associated deletion process.

At *CAN1* and at most of the individual hotspots examined, the combined effect of RNase H2 and high transcription was multiplicative (Fig. 2A, Fig. 3A–C). This is the result expected if transcription only enhances the recruitment of Top1 and the amount of rNMPs in DNA remains the same regardless of the level of transcription. At the (AG)₄ hotspot, however, we observed a greater-than-multiplicative effect when transcription and RNase H2 loss was combined. An intriguing possibility suggested by this result is that more rNMPs are incorporated into highly transcribed DNA. It is possible, for example, that the DNA damage associated with high-transcription is repaired outside of S phase, when the ratio of rNTPs to dNTPs is highest. Such a scenario could also explain why more dUTP is incorporated into highly transcribed DNA [7].

Previous studies have suggested that the most of the rNMPs in the yeast genome are incorporated by Pol ϵ rather than by Pol δ [15]. Deletions at (AG)₄ and (TC)₃ hotspots decreased in a *pol2-M644L rnh201 Δ* background (Fig. 4A), indicating that most rNMP-associated deletions at these positions are initiated by Top1 incision at rNMPs inserted specifically by Pol ϵ . The residual rNMP-dependent deletions observed in the *pol2-M644L rnh201 Δ* background could originate from rNMPs incorporated by the mutant Pol ϵ , WT Pol δ or WT Pol α . An alternative possibility is suggested by a recent *Escherichia coli* study, which demonstrated the incorporation of rNMPs into DNA by the Pol V translesion synthesis (TLS) DNA polymerase [30]. rNMP incorporation by the yeast Rev1, Pol ζ , and Pol η TLS polymerases and their potential contribution to rNMP-associated deletions has not yet been investigated.

Recapitulating what was seen under low-transcription conditions at the (AT)₂ hotspot [14], there was no significant rate increase in 2-bp deletions associated with RNase H2 loss under high-transcription conditions. The large increase in deletions associated with Top1-T722A expression (Fig. 4B) supports the model that deletions at this rNMP-independent hotspot originate from a trapped Top1cc [8, 14]. The source of Top1cc trapping, however, is unclear. In principle, any condition that perturbs the position of the 5'-OH, which is required for Top1-mediated re-ligation, could promote Top1cc trapping. Top1 incision near mismatches or transcription-associated damage, for example abasic sites, can prevent re-ligation *in vitro* [31]. In addition, collision between the RNA polymerase complex and a Top1cc can displace the 5'-OH and trap the enzyme [32]. In the latter case, the relevant Top1 cleavage site would be expected to be located on the transcribed strand.

Whereas we expected no effect of Top1-T722A expression on deletions at rNMP-associated hotspots, Top1-T722A expression actually decreased deletions at the (AG)₄ and (TC)₃ hotspots (Fig. 4B). There are two possible explanations for this result. First, more Top1-T722A is expected to be bound to DNA than WT enzyme, and this might limit the amount of free enzyme available to incise at rNMPs. A second possibility is that proficient re-ligation is required for the completion of the rNMP-associated deletion process. Although we cannot exclude the first explanation, we favor the latter because it is consistent with *in vitro* studies using purified Top1. When a Top1 cleaves less than 6 nucleotides upstream (5') of DNA end, the short fragment downstream of the transient Top1cc diffuses away, trapping the enzyme at the cleavage site – a classic “suicide” reaction [33] [34]. When a trapped Top1cc is provided with a single-strand DNA with 5'-OH, however, the enzyme can catalyze re-ligation [35]. The re-ligation by trapped Top1cc can occur across gaps, but it works better when the 5'-OH is perfectly aligned to the region immediately downstream (3') of the Top1cc [35]. Based on these *in vitro* data, we propose a sequential Top1 cleavage model for the rNMP-associated deletion process *in vivo* (Fig. 6). First, Top1 incises at an rNMP to generate a nick flanked by a 2',3'-cyclic phosphate and 5'-OH (step (i)). Either the released Top1 or a different Top1 then cleaves immediately upstream (5') of the nick, releasing a short oligonucleotide between the first and second Top1 cleavage sites and

trapping the enzyme (step (ii)–(iii)). Re-alignment of complementary strands facilitated by the tandem repeats directly juxtaposes the Top1cc and the 5′-OH (step (iv)). Using this 5′-OH, the Top1cc then catalyzes re-ligation (step (v)). The last step can explain why rNMP-associated deletions require proficient Top1 re-ligation activity and are reduced in the presence of the Top1-T722A protein. This model also can explain why the maximum deletion size observed for rNMP-dependent events thus far is 5 bp. For efficient diffusional loss of the short fragment between two cleavage sites, *in vitro* data suggest the fragment should be smaller than 6 nucleotides [36]. Finally, biochemical studies using calf thymus Top1 and vaccinia Top1 have shown that Top1 incision one nucleotide upstream of a nick is much less efficient than incision further away [37], which can explain the rarity of 1-bp relative to 2-bp deletions in the 6A run.

In summary, data reported here support the model that Top1-dependent deletions are initiated either by Top1cc trapping or by Top1 incision at an rNMP in DNA. We propose a new sequential-cleavage mechanism for the Top1-dependent, rNMP-associated deletions that are stimulated upon RNase H2 loss. RNase H2 is an essential enzyme in mice [38]; in humans, mutations in RNase H2 are associated with the neuroinflammatory disorder Aicardi-Goutières syndrome [39]. Either permanent single-strand break formation by Top1 incision at rNMPs, or a mutagenic effect at rNMPs may contribute to the deleterious consequences of RNase H2 loss in mammals. Alternatively, a Top1-dependent rNMP-associated deletion process might provide an alternative pathway for removing potentially deleterious rNMPs from genome. Finally, the Top1-dependent, rNMP-independent deletion process presumably reflects the removal of trapped Top1cc. The mode of action of chemotherapeutic agents such as camptothecin involve trapping of the Top1cc [13]. Top1-dependent mutagenesis, therefore, might be relevant to how cancer cells escape drug treatment and/or evolve to form secondary tumors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Thomas Kunkel for providing us with a plasmid to introduce the *pol2-M644L* allele [17]. We also thank Dr. Rodney Rothstein for Top1-T722A and WT Top1 plasmids [27]. We especially thank Dr. Stewart Shuman for suggesting the sequential Top1 cleavage model. This work was supported by grant 1R01GM101690 to SJR from NIH.

Abbreviations

TAM	transcription-associated mutagenesis
rNMP	ribonucleoside monophosphate
Top1	topoisomerase 1
Top1cc	topoisomerase 1 cleavage complex

References

1. Kim N, Jinks-Robertson S. Transcription as a source of genome instability. *Nat Rev Genet.* 2012; 13:204–214. [PubMed: 22330764]
2. Kim N, Abdulovic AL, Gealy R, Lippert MJ, Jinks-Robertson S. Transcription-associated mutagenesis in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication. *DNA Repair.* 2007; 6:1285–1296. [PubMed: 17398168]

3. Datta A, Jinks-Robertson S. Association of increased spontaneous mutation rates with high levels of transcription in yeast. *Science*. 1995; 268:1616–1619. [PubMed: 7777859]
4. Morey NJ, Greene CN, Jinks-Robertson S. Genetic analysis of transcription-associated mutation in *Saccharomyces cerevisiae*. *Genetics*. 2000; 154:109–120. [PubMed: 10628973]
5. Aguilera A, Garcia-Muse T. R loops: from transcription byproducts to threats to genome stability. *Mol Cell*. 2012; 46:115–124. [PubMed: 22541554]
6. Frederico LA, Kunkel TA, Shaw BR. A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry*. 1990; 29:2532–2537. [PubMed: 2185829]
7. Kim N, Jinks-Robertson S. dUTP incorporation into genomic DNA is linked to transcription in yeast. *Nature*. 2009; 459:1150–1153. [PubMed: 19448611]
8. Lippert MJ, Kim N, Cho JE, Larson RP, Schoenly NE, O'Shea SH, Jinks-Robertson S. Role for topoisomerase I in transcription-associated mutagenesis in yeast. *Proc Natl Acad Sci USA*. 2011; 108:698–703. [PubMed: 21177427]
9. Takahashi T, Burguiere-Slezak G, Van der Kemp PA, Boiteux S. Topoisomerase I provokes the formation of short deletions in repeated sequences upon high transcription in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA*. 2011; 108:692–697. [PubMed: 21177431]
10. Wu HY, Shyy SH, Wang JC, Liu LF. Transcription generates positively and negatively supercoiled domains in the template. *Cell*. 1988; 53:433–440. [PubMed: 2835168]
11. Wang JC. Cellular roles of DNA topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol*. 2002; 3:430–440. [PubMed: 12042765]
12. Phatnani HP, Jones JC, Greenleaf AL. Expanding the functional repertoire of CTD kinase I and RNA polymerase II: novel phosphoCTD-associating proteins in the yeast proteome. *Biochemistry*. 2004; 43:15702–15719. [PubMed: 15595826]
13. Pommier Y, Barcelo JM, Rao VA, Sordet O, Jobson AG, Thibaut L, Miao ZH, Seiler JA, Zhang H, Marchand C, Agama K, Nitiss JL, Redon C. Repair of topoisomerase I-mediated DNA damage. *Prog Nucleic Acid Res Mol Biol*. 2006; 81:179–229. [PubMed: 16891172]
14. Kim N, Huang SN, Williams JS, Li YC, Clark AB, Cho JE, Kunkel TA, Pommier Y, Jinks-Robertson S. Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. *Science*. 2011; 332:1561–1564. [PubMed: 21700875]
15. Nick McElhinny SA, Watts BE, Kumar D, Watt DL, Lundstrom EB, Burgers PM, Johansson E, Chabes A, Kunkel TA. Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc Natl Acad Sci USA*. 2010; 107:4949–4954. [PubMed: 20194773]
16. Rydberg B, Game J. Excision of misincorporated ribonucleotides in DNA by RNase H (type 2) and FEN-1 in cell-free extracts. *Proc Natl Acad Sci USA*. 2002; 99:16654–16659. [PubMed: 12475934]
17. Nick McElhinny SA, Kumar D, Clark AB, Watt DL, Watts BE, Lundstrom EB, Johansson E, Chabes A, Kunkel TA. Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol*. 2010; 6:774–781. [PubMed: 20729855]
18. Sparks JL, Chon H, Cerritelli SM, Kunkel TA, Johansson E, Crouch RJ, Burgers PM. RNase H2-Initiated Ribonucleotide Excision Repair. *Mol Cell*. 2012; 47:980–986. [PubMed: 22864116]
19. Clark AB, Lujan SA, Kissling GE, Kunkel TA. Mismatch repair-independent tandem repeat sequence instability resulting from ribonucleotide incorporation by DNA polymerase epsilon. *DNA Repair*. 2011; 10:476–482. [PubMed: 21414850]
20. Sekiguchi J, Shuman S. Site-specific ribonuclease activity of eukaryotic DNA topoisomerase I. *Mol Cell*. 1997; 1:89–97. [PubMed: 9659906]
21. Sikorski RS, Hieter P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics*. 1989; 122:19–27. [PubMed: 2659436]
22. Lea DE, Coulson CA. The distribution of the numbers of mutants in bacterial populations. *J Genet*. 1949; 49:264–285.
23. Spell RM, Jinks-Robertson S. Determination of mitotic recombination rates by fluctuation analysis in *Saccharomyces cerevisiae*. *Methods Mol Biol*. 2004; 262:3–12. [PubMed: 14769952]
24. Altman, DG. *Practical Statistics for Medical Research*. Chapman & Hall; London: 1991.

25. Lehner K, Jinks-Robertson S. The mismatch repair system promotes DNA polymerase zeta-dependent translesion synthesis in yeast. *Proc Natl Acad Sci USA*. 2009; 106:5749–5754. [PubMed: 19307574]
26. Megonigal MD, Fertala J, Bjornsti MA. Alterations in the catalytic activity of yeast DNA topoisomerase I result in cell cycle arrest and cell death. *J Biol Chem*. 1997; 272:12801–12808. [PubMed: 9139740]
27. Reid RJ, Gonzalez-Barrera S, Sunjevaric I, Alvaro D, Ciccone S, Wagner M, Rothstein R. Selective ploidy ablation, a high-throughput plasmid transfer protocol, identifies new genes affecting topoisomerase I-induced DNA damage. *Genome Res*. 2011; 21:477–486. [PubMed: 21173034]
28. Greene CN, Jinks-Robertson S. Frameshift intermediates in homopolymer runs are removed efficiently by yeast mismatch repair proteins. *Mol Cell Biol*. 1997; 17:2844–2850. [PubMed: 9111356]
29. Harfe BD, Jinks-Robertson S. Removal of frameshift intermediates by mismatch repair proteins in *Saccharomyces cerevisiae*. *Mol Cell Biol*. 1999; 19:4766–4773. [PubMed: 10373526]
30. Kuban W, Vaisman A, McDonald JP, Karata K, Yang W, Goodman MF, Woodgate R. *Escherichia coli* UmuC active site mutants: Effects on translesion DNA synthesis, mutagenesis and cell survival. *DNA Repair*. 2012; 11:726–732. [PubMed: 22784977]
31. Pourquier P, Pilon AA, Kohlhagen G, Mazumder A, Sharma A, Pommier Y. Trapping of mammalian topoisomerase I and recombinations induced by damaged DNA containing nicks or gaps. Importance of DNA end phosphorylation and camptothecin effects. *J Biol Chem*. 1997; 272:26441–26447. [PubMed: 9334220]
32. Wu J, Liu LF. Processing of topoisomerase I cleavable complexes into DNA damage by transcription. *Nucleic Acids Res*. 1997; 25:4181–4186. [PubMed: 9336444]
33. Shuman S, Prescott J. Specific DNA cleavage and binding by vaccinia virus DNA topoisomerase I. *J Biol Chem*. 1990; 265:17826–17836. [PubMed: 2170398]
34. Christiansen K, Westergaard O. Mapping of eukaryotic DNA topoisomerase I catalyzed cleavage without concomitant religation in the vicinity of DNA structural anomalies. *Biochim Biophys Acta*. 1999; 1489:249–262. [PubMed: 10673027]
35. Henningfeld KA, Hecht SM. A model for topoisomerase I-mediated insertions and deletions with duplex DNA substrates containing branches, nicks, and gaps. *Biochemistry*. 1995; 34:6120–6129. [PubMed: 7742316]
36. Shuman S. Site-specific interaction of vaccinia virus topoisomerase I with duplex DNA. Minimal DNA substrate for strand cleavage *in vitro*. *J Biol Chem*. 1991; 266:20576–20577. [PubMed: 1682312]
37. Cheng C, Shuman S. Site-specific DNA transesterification by vaccinia topoisomerase: role of specific phosphates and nucleosides. *Biochemistry*. 1999; 38:16599–16612. [PubMed: 10600122]
38. Reijns MA, Rabe B, Rigby RE, Mill P, Astell KR, Lettice LA, Boyle S, Leitch A, Keighren M, Kilanowski F, Devenney PS, Sexton D, Grimes G, Holt IJ, Hill RE, Taylor MS, Lawson KA, Dorin JR, Jackson AP. Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell*. 2012; 149:1008–1022. [PubMed: 22579044]
39. Crow YJ, Leitch A, Hayward BE, Garner A, Parmar R, Griffith E, Ali M, Semple C, Aicardi J, Babul-Hirji R, Baumann C, Baxter P, Bertini E, Chandler KE, Chitayat D, Cau D, Dery C, Fazzi E, Goizet C, King MD, Klepper J, Lacombe D, Lanzi G, Lyall H, Martinez-Frias ML, Mathieu M, McKeown C, Monier A, Oade Y, Quarrell OW, Rittey CD, Rogers RC, Sanchis A, Stephenson JB, Tacke U, Till M, Tolmie JL, Tomlin P, Voit T, Weschke B, Woods CG, Lebon P, Bonthron DT, Ponting CP, Jackson AP. Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutières syndrome and mimic congenital viral brain infection. *Nat Genet*. 2006; 38:910–916. [PubMed: 16845400]

Highlights

- Top1-dependent rNMP-independent deletions result from trapped Top1cc.
- Sequential-cleavage mechanism for Top 1-and rNMP-dependent deletions is proposed.
- Transcription and rNMPs have a multiplicative effect on Top1-dependent deletions.

Fig. 1. End-processing model for Top1-dependent deletions

Light grey boxes indicate the relevant dinucleotide repeat. Yellow arrows designate Top1 cleavage. See section 1 for details.

Fig. 2. 2–5 bp deletions at *CAN1*
A. Rate of 2–5 bp deletions. The rate under low-transcription in an *rnh201Δ top1Δ* background is not shown because it was determined in a different strain background [14]. Error bars show 95% confidence intervals. B. Locations of 2–5 bp deletions. Only a part of *CAN1* ORF is shown. Grey bars over the sequence and red bars below the sequence represent deletion events observed in the WT and *rnh201Δ* background, respectively (Table S2). The length of the bars corresponds to the size of deletion. Light grey boxes highlight all dinucleotide repeats. The (AT)₂, (TC)₃, and (AG)₄ hotspot-containing sequences transplanted into the *lys2ΔA746,NR* reversion window are highlighted with yellow boxes.

Fig. 3. Reversion rates of *lys2ΔA746NR* alleles

The *LYS2* and *TET* promoters were used for low- and high-transcription conditions, respectively. Grey bars represent 2-bp deletions (Table S2). White bars represent other events. Low-transcription data and WT high-transcription data were previously reported [8, 14]. The 95% confidence intervals are indicated.

Fig. 4. Effect of Pol2-M644L and Top1-T722A expression on 2-bp deletions under high-transcription conditions

A. Effect of Pol2-M644L on 2-bp deletion rates in an *rnh201Δ* background (Table S2). B. Effect of Top1-T722A expression on 2-bp deletion frequencies in *top1Δ rnh201Δ* backgrounds (Table S3). Frequencies observed with an empty vector control or WT Top1 expression also are shown. Error bars represent 95% confidence intervals of rates (A) and median frequencies (B).

Fig. 5. Reversion rates of *lys2ΔBgl* and *lys2ΔA746* alleles

A. The *lys2ΔBgl* allele was under control of the *TET* promoter. Low- and high-transcription conditions were created by growth in the presence or absence of doxycycline, respectively. Dark grey bars correspond to 4-bp deletions at (AGCT)₂ and light grey bars to 1-bp deletions at the 6A run (Table S4). B. The *LYS2* and *TET* promoters were used for low- and high- transcription conditions, respectively. Black bars correspond to 2-bp deletions at the 6A run and white bars to other events (Table S4). The 95% confidence interval of the total Lys⁺ rate is indicated.

Fig. 6. Sequential Top1 cleavage model for rNMP-associated deletions

Light grey boxes indicate the relevant dinucleotide repeat. Yellow arrows designate Top1 cleavage. See section 4 for details.