Published in final edited form as:
Science. 2011 June 24; 332(6037): 1561–1564. doi:10.1126/science.1205016.

Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase 1

Nayun Kim[1], **Sharyin N. Huang**[2], **Jessica Williams**[3], **Yue C. Li**[1], **Alan B. Clark**[3], **Jang-Eun Cho**[1], **Thomas A. Kunkel**[3], **Yves Pommier**[2], and **Sue Jinks-Robertson**[1]

[1]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710

[2]Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892

[3]Laboratory of Molecular Genetics and Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709

Abstract

The RNase H class of enzymes degrades the RNA component of RNA:DNA hybrids and is important in nucleic acid metabolism. RNase H2 is specialized to remove single ribonucleotides (rNMPs) from duplex DNA, and its absence in budding yeast has been associated with the accumulation of deletions within short tandem repeats. Here, we demonstrate that rNMP-associated deletion formation requires the activity of Top1, a topoisomerase that relaxes supercoils by reversibly nicking duplex DNA. The reported studies extend the role of Top1 to include the processing of rNMPs in genomic DNA into irreversible single-strand breaks, an activity that can have distinct mutagenic consequences and may be relevant to human disease.

The exclusion and removal of rNMPs from DNA are important for the stability and function of the genome. In *Saccharomyces cerevisiae*, the introduction of an rNMP-permissive form of DNA polymerase ε into a strain lacking RNase H2 confers a mutator phenotype and is associated with the accumulation of a distinct mutation class: deletions within short (2–5 bp) tandem repeats (1). In contrast to similar mutations initiated by DNA polymerase slippage during genome replication, however, the rNMP-associated deletion intermediates are not substrates for the postreplicative mismatch repair machinery (2). A similar deletion signature is associated with high levels of transcription in yeast and requires the activity of Top1 (3, 4), a type 1B topoisomerase important for removing transcription-associated supercoils (5). Here, we demonstrate that rNMP-associated deletions are likewise dependent on Top1 activity, map *in vitro* the positions of Top1 cleavage at deletion hotspots identified *in vivo*, and confirm that Top1 has endoribonuclease activity when an rNMP is substituted at the scissile phosphate.

The *CAN1* gene encodes arginine permease, the loss of which confers resistance to the toxic arginine analog canavanine (Can-R phenotype). To determine the effect of persistent rNMPs on *CAN1* mutagenesis in yeast, we deleted the *RNH201* gene, which encodes the catalytic subunit of RNase H2 (6). While there was only a small elevation in the Can-R rate in the *rnh201* background, there was a substantial change in the corresponding mutation spectrum, with ~40% of mutations being deletions of 2–5 bp (partial and complete spectra are presented in Fig. 1 and Fig. S1, respectively). To examine whether Top1 activity is relevant to rNMP-associated mutagenesis, we deleted the *TOP1* gene from the *rnh201* background. The rate of short deletions in the double mutant reverted to that observed in the WT strain, demonstrating that Top1 is required for the rNMP-associated deletion signature. Hereafter, we focus on the 2-bp deletion class, which localizes to discrete hotspots that coincide with

short dinucleotide repeats. Some of these hotspots appear to be unique to the *rnh201* background while others are Top1-dependent hotspots previously observed under high-transcription conditions [e.g., the $(AG)_4$ and $(TC)_3$ runs at nt 254 and 1448 of *CAN1*, respectively] (3).

Small (20–30 bp) fragments containing representative hotspots identified in the *CAN1* reporter were transplanted into the *lys2ΔA746NR* frameshift reversion assay, which detects net 1-bp insertions (7). The $(AG)_4$ and $(TC)_3$ hotspots were examined in this context, as well as the transcription-dependent $(AT)_2$ hotspot at position 1127 (3). A substantial proportion (21%) of *lys2ΔA746NR*,$(AG)_4$ revertants contained a 2-bp deletion at the introduced hotspot (gray bars in Fig. 2; see Table S1 and Fig. S2 for corresponding rates and spectra, respectively). Recapitulating what was seen in the *CAN1* assay, deletion of *RNH201* had the greatest effect on the $(AG)_4$ hotspot, elevating the Lys^+ rate ~60 fold and shifting the spectrum towards events at the introduced hotspot (from 21% to 95%). In the *rnh201* background, the reversion rate of the *lys2ΔA746NR,*$(TC)_3$ allele was elevated 7-fold, and the proportion of mutations at the hotspot increased from 51% to 89%. At both the $(AG)_4$ and $(TC)_3$ hotspots, the 2-bp deletions were absent in the *rnh201 top1* double mutant, confirming their dependence on Top1. In contrast to the $(AG)_4$ and $(TC)_3$ hotspots, loss of Rnh201 had little, if any, effect on activity of the Top1-dependent $(AT)_2$ hotspot. We suggest either that rNMPs are not incorporated as often near the $(AT)_2$ hotspot, or that rNMPs incorporated here are refractory to RNase H2 activity.

Loss of Rnh201 greatly elevates 4-bp deletions within a 4-bp tandem repeat in the *lys2ΔBgl* assay (8), an assay that queries the same sequence as the *lys2ΔA746NR* system, but detects net 1-bp deletions (9). Although it was suggested that the 4-bp deletions reflect defective Okazaki-fragment processing, their molecular similarity to 2-bp deletions in the *CAN1* assay prompted us to examine their dependence on Top1. Whereas the 4-bp deletions comprised 69% of the spectrum in an *rnh201* mutant, they were completely eliminated upon additional deletion of *TOP1* (Fig. 2, Table S1 and Fig. S2D).

The Top1-dependent deletions identified under high-transcription conditions were proposed to reflect repair of trapped Top1 cleavage complexes (3, 4). In such complexes, the enzyme is linked to DNA via a 3′-phosphotyrosyl bond, leaving a 5′-OH on the other side of the nick (Fig. 3A). To explain the observed deletion pattern, it was suggested that the Top1-generated ends are processed into a gap corresponding in the size to that of the ensuing deletion (Fig. 4). The significance of the tandem repeat is that it can stabilize misalignment between complementary DNA strands, thereby bringing the ends flanking the gap together and facilitating their ligation. With regard to the Top1-dependent deletions observed here in the absence of RNase H2, biochemical studies have demonstrated that Top1 can act as an endonuclease if an rNMP is present at the cleavage site (10). In this reaction, the 2′ OH of the ribose attacks the 3′-phosphotyrosyl linkage between the enzyme and ribonucleotide, releasing Top1 and leaving a 2′,3′ cyclic phosphate end (Fig. 3B). In a manner analogous to that described above for a trapped cleavage complex, we suggest that subsequent processing of the cyclic end could generate a gap within the relevant tandem repeat, followed by misalignment between complementary strands to facilitate ligation (Fig. 4).

The molecular scenarios outlined above require Top1 cleavage either within or immediately adjacent to the tandem repeats where deletions occur. The small size and degeneracy of the Top1 consensus sequence, however, make it difficult to predict cleavage sites (11, 12), and mapping cleavage sites *in vitro* can be problematic because of the rapid religation of Top1-cleaved DNA. The Top1-specific drug camptothecin (CPT) is thus often used to stabilize the covalent cleavage intermediate, which can then be reversed in a time-dependent manner following an increase in salt concentration (13, 14). We examined CPT-stabilized Top1

cleavage intermediates produced when ~30 bp substrates containing the $(AT)_2$, $(TC)_3$ or $(AG)_4$ hotspot (Fig. 3 and Fig. S3) were radioactively labeled on one of the 3′ ends. Under these conditions, Top1 cleavage produces labeled fragments shorter than full-length DNA. For the $(AT)_2$ construct, the most prominent product was an 18-nt fragment ("f" in Fig. 3C), indicating a major Top1 cleavage site on the transcribed stand in the middle of the dinucleotide repeat. There were weaker sites within the same fragment, as is often seen in Top1 cleavage assays.

Substituting ribo-uracil (rU) for dT at the major cleavage site in the $(AT)_2$ fragment enhanced production of the corresponding Top1-generated fragment, presumably because the enzyme was released before the backbone could be religated (10). Inclusion of CPT enhanced Top1 cleavage complex formation at other, CPT-dependent sites, thus competing with and reducing cleavage at the rU-substituted site. While the addition of high salt reduced the intensity of the CPT-dependent fragments, the fragment corresponding to cleavage at rU did not change, indicating irreversible nicking at this specific site. To further examine the nature of the Top1 cleavage products, the corresponding transcribed strand was labeled at its 5′ end (Fig. 3C). With the dT-containing fragment, the 5′-labeled intermediate failed to enter the gel, consistent with covalent linkage of Top1 to the 3′ end at the cleavage site. With the rU substrate, however, a 5′-labeled, 12-nt cleavage product with a terminal phosphate was generated, indicating release of Top1 following cleavage, as predicted ("f" in bottom panel of Fig. 3C).

Results of similar analyses with an ~30 bp duplex fragment containing the $(TC)_3$ hotspot are presented in Fig. 3D and Fig. S3. With the $(TC)_3$ hotspot, a single CPT-stabilized, salt-reversible Top1 cleavage product was detected when the transcribed strand was labeled at the 3′ end (fragment "h"); no cleavage product was seen when the 5′ end was labeled. Substitution of rU for the relevant dT residue yielded an irreversible Top1 cleavage product of appropriate size when either the 3′ or 5′ end of the transcribed strand was labeled. The sole Top1 cleavage site detected in the $(TC)_3$ fragment is immediately adjacent to, but is not within, the dinucleotide repeat. This has potential relevance to requisite end-processing step(s) as well as to the mechanism of deletion formation *in vivo*. Finally, we examined possible positions of Top1 cleavage in a fragment containing the $(AG)_4$ hotspot. In this case, two cleavage sites were detected on each strand (Fig. S3). We speculate that the site immediately adjacent to the $(AG)_4$ repeat on the transcribed strand is the one most likely relevant to Top1-dependent mutagenesis.

Here we have shown that, in addition to its well-known role in relaxing DNA supercoils, Top1 can initiate the removal of rNMPs from yeast genomic DNA. Because the intermediates formed during mutagenesis associated with rNMP incorporation by DNA polymerase ε are not substrates for mismatch repair (2), they likely arise outside the context of normal DNA replication and potentially could be a source of mutagenesis in terminally differentiated cells. Elucidating the pathway(s) that resolve the cyclic end produced when Top1 incises at an rNMP is clearly of importance, as resolution/repair can be highly mutagenic in yeast and is expected to be similarly mutagenic in other organisms. There are a large number of proteins/complexes that have been implicated in the processing of CPT-stabilized Top1 intermediates that become trapped during DNA replication (15), some of which may be relevant to the processing we predict occurs at Top1-generated nicks. Alterations in RNase H2 are one cause of the severe autoimmune disease Acardi Goutiéres Syndrome (16), which is characterized by neurological dysfunction similar to that associated with congenital viral infection. An intriguing possibility is that the RNase H2-like activity of Top1 might have relevance to disease etiology. Finally, the mutagenic activity of Top1 identified here could be relevant to cases of mismatch repair-independent microsatellite instability described in human tumor cells (17).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

S.J.-R. was supported by National Institutes of Health grants R01 GM38464 and R01 GM93197; Y.P. was supported by the Intramural Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health; and T.A.K. was supported by Projects Z01 ES065070 and Z01 ES065089 from the Division of Intramural Research of the National Institutes of Health, National Institute of Environmental Health Sciences.

References

1. McElhinny SA, et al. Nat. Chem. Biol. 2010
2. Clark AB, Lujan SA, Kissling GE, Kunkel TA. DNA Rep. 2011 **in press**.
3. Lippert MJ, et al. Proc. Natl. Acad. Sci. USA. 2011; 108:698. [PubMed: 21177427]
4. Takahashi T, Burguiere-Slezak G, Van der Kemp PA, Boiteux S. Proc. Natl. Acad. Sci. USA. 2011; 108:692. [PubMed: 21177431]
5. Wang JC. Nat. Rev. Mol. Cell. Biol. 2002; 3:430. [PubMed: 12042765]
6. Jeong HS, Backlund PS, Chen HC, Karavanov AA, Crouch RJ. Nucleic Acids Res. 2004; 32:407. [PubMed: 14734815]
7. Harfe BD, Jinks-Robertson S. Mol. Cell. Biol. 1999; 19:4766. [PubMed: 10373526]
8. Chen JZ, Qiu J, Shen B, Holmquist GP. Nucleic Acids Res. 2000; 28:3649. [PubMed: 10982888]
9. Greene CN, Jinks-Robertson S. Mol. Cell. Biol. 1997; 17:2844. [PubMed: 9111356]
10. Sekiguchi J, Shuman S. Mol. Cell. 1997; 1:89. [PubMed: 9659906]
11. Porter SE, Champoux JJ. Mol. Cell. Biol. 1989; 9:541. [PubMed: 2540421]
12. Jaxel C, Capranico G, Kerrigan D, Kohn KW, Pommier Y. J. Biol. Chem. 1991; 266:20418. [PubMed: 1657924]
13. Hsiang YH, Liu LF. Cancer Res. 1988; 48:1722. [PubMed: 2832051]
14. Pommier Y, Leo E, Zhang H, Marchand C. Chem. Biol. 2010; 17:421. [PubMed: 20534341]
15. Pommier Y, et al. Prog. Nucleic Acid Res. Mol. Biol. 2006; 81:179. [PubMed: 16891172]
16. Crow YJ, et al. Nat. Genet. 2006; 38:910. [PubMed: 16845400]
17. Clark AB, Kunkel TA. Cell Cycle. 2010; 9:4422. [PubMed: 21088478]
18. Pavlov YI, et al. Proc. Natl. Acad. Sci. USA. 2002; 99:9954. [PubMed: 12119399]
19. Goldstein AL, McCusker JH. Yeast. 1999; 15:1541. [PubMed: 10514571]
20. Shcherbakova PV, Kunkel TA. Mol. Cell. Biol. 1999; 19:3177. [PubMed: 10082584]
21. Sikorski RS, Hieter P. Genetics. 1989; 122:19. [PubMed: 2659436]
22. Kim N, Abdulovic AL, Gealy R, Lippert MJ, Jinks-Robertson S. DNA Repair. 2007; 6:1285. [PubMed: 17398168]
23. Dexheimer TS, Pommier Y. Nat. Protoc. 2008; 3:1736. [PubMed: 18927559]
24. Pourquier P, et al. J. Biol. Chem. 1999; 274:8516. [PubMed: 10085084]

A *rnh201Δ* (N=185)

B *rnh201Δ top1Δ* (N=165)

C

Strain	Mutation rate (x 10^{-8})				
	BS	1-bp indel	2-5 bp del	Other	Total
WT (N=82)	8.3	0.9	0.2	0.6	10 (7.7-13)
rnh201Δ (N=185)	10.4	4.2	10.4	1.0	26 (22-32)
rnh201Δ top1Δ (N=165)	12.4	4.5	0.2	0.9	18 (17-22)

Figure 1. Mutagenesis in the *CAN1* forward mutation assay

(A and **B)** show partial mutation spectra of Can-R mutants (complete spectra are in Figure S1). Nucleotides are numbered beginning with the ATG start codon. Base substitutions and indels are in red below and above the sequence, respectively; lengths of red bars correspond to deletion sizes. Sequences transplanted into the *lys2ΔA746NR* assay are highlighted in yellow. **C** presents rates of individual mutation types at *CAN1*. N, number of mutants sequenced; indel, insertion/deletion; BS, base substitution. 95% confidence intervals are in parentheses below total rates.

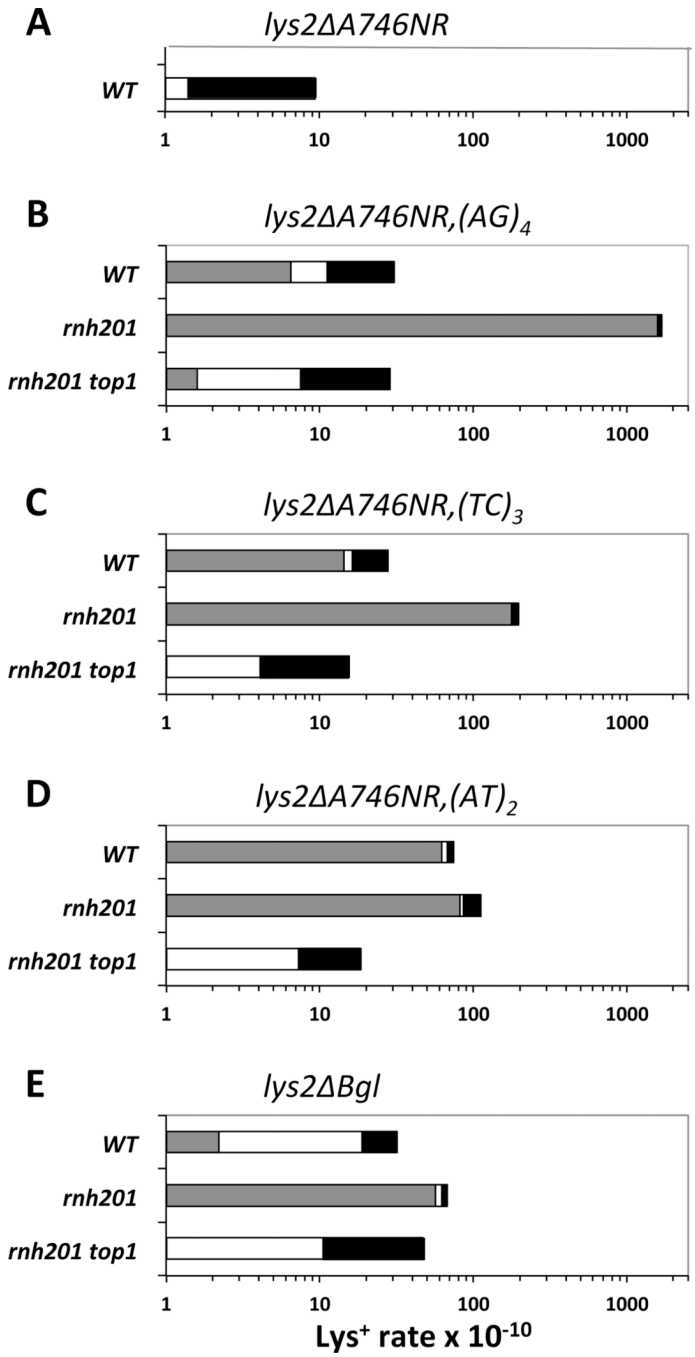

Figure 2. Rates of mutation types in *lys2* frameshift reversion assays

Gray bars correspond to 2-bp deletions at the introduced hotspot. White and black bars represent 1-bp indels and all other classes of mutations, respectively.

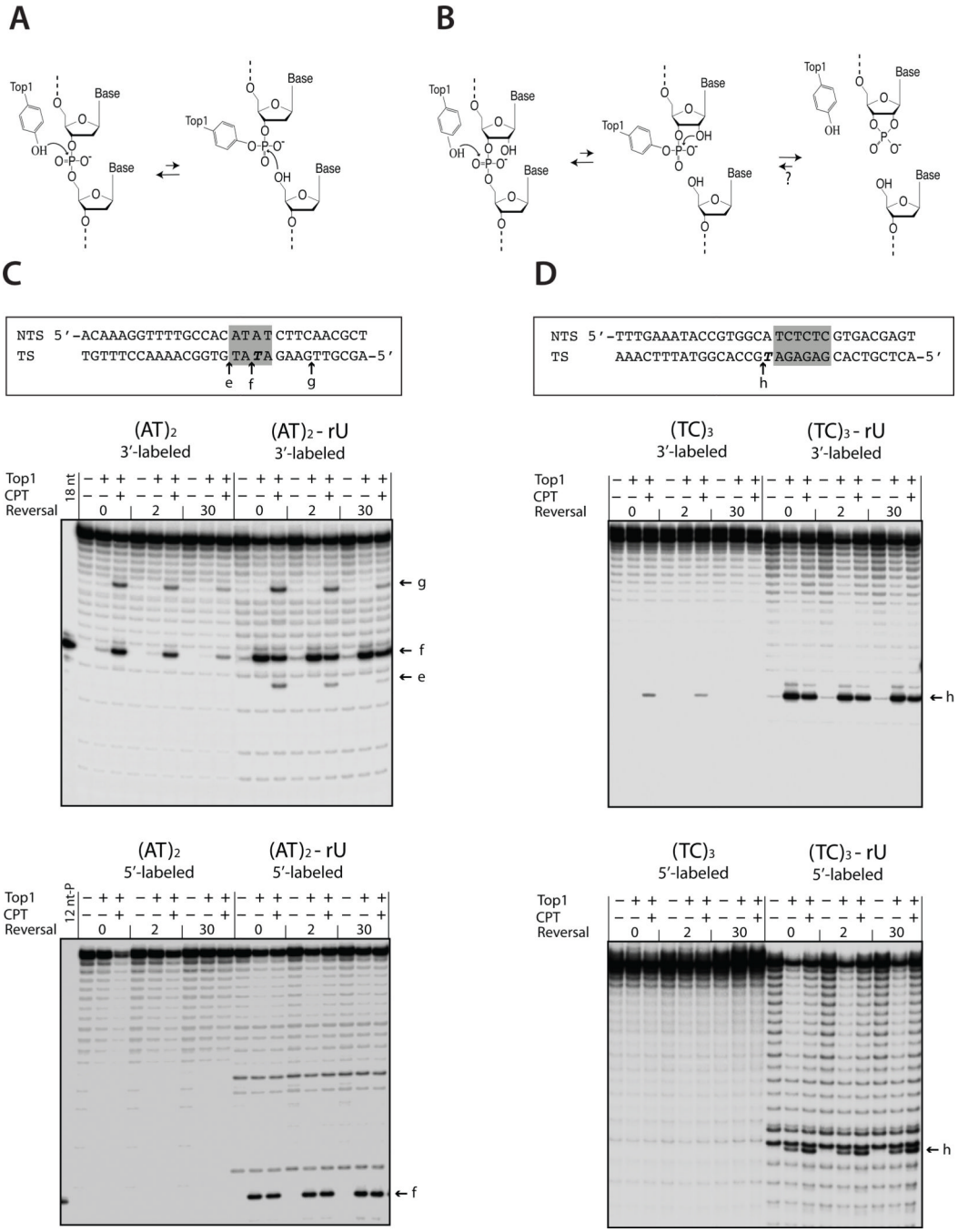

Figure 3. Top1 cleavage assays

(A and **B)** illustrate the mechanism of Top1 cleavage at a scissile dNMP versus rNMP, respectively. **(C** and **D)** show cleavage of $(AT)_2$ and $(TC)_3$ hotspot fragments, respectively, by human Top1. In the fragment sequences, relevant dinucleotide repeats are highlighted in gray and ribo-substituted nucleotides are in bold italics. Transcribed and nontranscribed strands are designated TS and NTS, respectively; the TS strand was radioactively labeled at either the 3′ or 5′ end as indicated. Labeled arrows indicate positions of Top1 cleavage. Time points for the salt reversal experiments are in min.

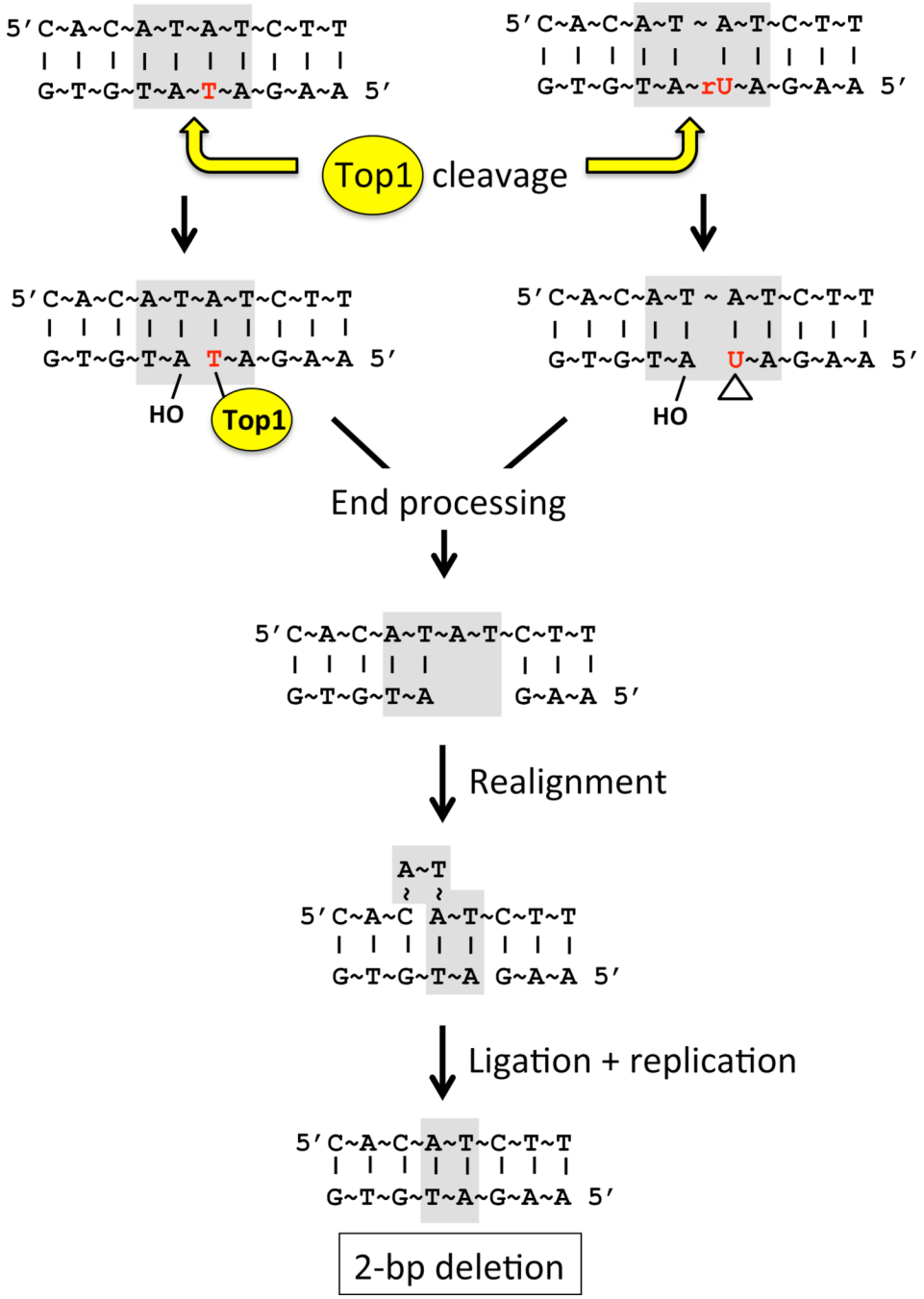

Figure 4. Model for Top1-initiated deletions

The $(AT)_2$ hotspot sequence is shown; the top strand is the nontranscribed strand and the dinucleotide repeat highlighted in gray. The cyclic 2′,3′ phosphate formed by Top1 cleavage at rUMP is indicated by a triangle.