



**Reverse transcriptase model of somatic hypermutation: Predicted molecular steps in transcription-coupled DNA and RNA deamination and reverse transcription.** See publication number 23 for discussion and references. Shown are some key hypothesised intermediates highlighted for the generation of the prominent strand biased mutations observed in SHM: A-to-G, G-to-A and G-to-C signatures. Black lines are DNA strands, red lines are mRNA, blue lines are cDNA strands copied off mRNA. Steps **A.** through to **D.** show various mutated DNA and RNA intermediates and substrate complexes for both deamination reactions and cDNA synthesis. In brief, mutations introduced first at the DNA level (AID-mediated C-to-U or abasic sites in the TS) then further modified on copying into mRNA by RNA Pol II (Kuraoka et al., 2003 *J Biol Chem* **278**: 7294-7299) and adenosine-to-inosine (A-to-I) RNA editing (Ref 21) followed by copying of the mutated mRNA by error-prone reverse transcription via Pol- $\eta$  (Ref 16) which targets A template sites in duplex mRNA in the context of a WA motif (Ref 21). Then follows strand invasion and integration of newly synthesised cDNA TS. In more detail: **A.** RNA Pol II introduces mutations in mRNA as it copies the AID-mediated lesions in TS DNA (Kuraoka et al., 2003), followed by ADAR1-mediated A-to-I RNA editing of WA sites in dsRNA stem(-loops) forming in nascent mRNA near the transcription bubble (Ref 21 and Fig 1). **B.** Formation of reverse transcriptase -priming substrate by annealing of nicked TS strand with an exposed 3'-OH end. This could arise due to excision at a previous AID-mediated abasic site or an excision introduced by an endonuclease activity associated with the MSH2-MSH6 heterodimer engaging a U:G mispaired lesion. **C.** Extension of new TS by cDNA synthesis from the 3'-OH end copying the already mutated mRNA template (with I base pairing preferentially, like G, with C). **D.** Then an unknown and indeterminate number of steps involving strand invasion, heteroduplex formation and/or resolution of heteroduplex, full length copying of newly synthesized transcribed strand cDNA.