Trinucleotide expansion in disease: why is there a length threshold?
Do-Yup Lee¹ and Cynthia T McMurray²

Trinucleotide repeats (TNRs) expansion disorders are severe neurodegenerative and neuromuscular disorders that arise from inheriting a long tract (30–50 copies) of a trinucleotide unit within or near an expressed gene (Figure 1a). The mutation is referred to as ‘trinucleotide expansion’ since the number of triplet units in a mutated gene is greater than the number found in the normal gene. Expansion becomes obvious once the number of repeating units passes a critical threshold length, but what happens at the threshold to render the repeating tract unstable? Here we discuss DNA-dependent and RNA-dependent models by which a particular DNA length permits a rapid transition to an unstable state.

Addresses
¹ Department of Bio and Fermentation Convergence Technology, Kookmin University, 136-702, Republic of Korea
² Life Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

Corresponding author: McMurray, Cynthia T. (ctmcmurray@lbl.gov)

Trinucleotide expansion is the underlying basis for disease toxicity in a number of severe hereditary diseases [1,2,3,4], and occurs both in the germ line and in somatic tissues with age. The general steps of expansion in simple terms are three: structure formation, heteroduplex resolution, and gap filling synthesis (Figure 1b). Over the past years, many reviews (including our own) have focused on the first step: how heteroduplex DNA structures form [1,4,5,6] (Figure 1c). Indeed, all data are consistent with a model in which heteroduplex structures are the basis for expansion, which arises broadly from classes of de novo excision repair, replication errors, and replication arrest and restart [1,4,5]. All of these mechanisms invoke their own machinery to carry out heteroduplex resolution, and distinct polymerases to complete gap-filling synthesis (Table 1). DNA expansion itself appears to be independent of position of the repeat tract, other than that it must reside in or around genes to cause observable abnormalities (Figure 1). But how do expansions begin? Here, we will consider one of the oldest questions and most puzzling feature of expansion: its length threshold.

What is an expansion threshold? Expansion observed in all TNR diseases requires a pre-existing long tract of TNRs units before there is a significant probability of instability (Figure 1a). Normal allele lengths are stable, and there is no ‘jumping’ from a normal to a disease tract length [7,8] (Figure 1a). Only when an allele is of critical copy number (the threshold) does expansion become probable within the lifetime of a human, and modulate a transition from pre-mutation to full-mutation length TNR tract [7–10]. The fact that expansion becomes probable only after a threshold length is reached suggests that expansion is strongly DNA-dependent, but why does tract length matter? In this review, we discuss three major models that provide possible explanations for a length threshold in light of recent findings: firstly length-dependent reannealing of DNA or DNA–RNA hybrids, secondly coding for a minimum length of RNA and protein sufficient to induce toxicity, and finally metabolism. These mechanisms are not mutually exclusive, but some are more likely than others.

Does the stability and size of heteroduplex loops in DNA govern the threshold?
One of the oldest and perhaps most intuitive explanation for a threshold is a minimal length at which a heteroduplex DNA intermediate becomes stable [1,2,3,4,5,6]. Indeed, we demonstrated as early as 1995 that triplet repeats formed hairpins with repeating units of two CG pairs and a mismatch, which explained their aberrant migration on gels [11]. At the same time, Wells and co-workers observed that instability occurred in bacteria by slippage [12]. However, a structural stability model for threshold is not entirely satisfying. Loop sizes of only a few repeats are thermodynamically stable in replication slippage reactions [6], and the MutL endonuclease that resolves small loops in DNA operates efficiently at 1–4 contiguous triplet units [13].

However, the sizes of the heteroduplex loops that occur during repair are expected to be larger. The excision patch of transcription coupled repair (TCR) and nucleotide excision repair (NER) is typically around 15–20 bases [14], corresponding to a fold-back structure of 5–7 repeats. Strand displacement during long patch BER is...
Features of expansion and its threshold. (a) Generic representation of threshold limits for some representative disease alleles from distinct TNR disorders. The inverted purple triangles represent the size ranges associated with the normal, threshold, and disease length TNR alleles. In white is the threshold length for representative TNRs: CGG in the 5′-untranslated (5′UT) region characterizes the FMR-1 gene; GAA in an intron (lines) characterizes Friedreich’s ataxia gene; CAG in a coding region (exon) characterizes the Huntington’s gene; CTG in the 3′-untranslated (3′UT) region characterizes the myotonic dystrophy 1 (DM1) gene. The threshold limit is also referred to as the premutation length, as all full mutations arise from lengths at the upper range of normal allele and the lower edge of disease allele lengths. Below that range are stable normal repeats, and above the ranges at which expansion exists. The premutation lengths as shown are approximate sizes since there is no precise range THD is threshold. (b) The three most basic steps of expansion. (c) Distinct types of heteroduplex DNA loops are proposed as precursors to expansion: hairpins (1); cruciform (2); quadruplex (3); H-DNA triplet helix (4). (c) is taken from [5*].

A kinetic model for the threshold on the DNA level is more likely. At any single strand break or on Okazaki fragments, free ends are in flux on and off DNA, and there is inherent competition between duplex reformation (no mutation) and structure formation at the frayed end (mutation intermediate). The threshold transition length may simply reflect the length at which the lifetime of self-pairing in heteroduplex DNA becomes long enough to exceed the rate of gap filling synthesis (which would prevent duplex reannealing). The resulting flap folds-back to initiate structure formation at the TNR sequence. Indeed, we tested at least part of this idea by following duplex reannealing of complementary hairpins of 10 (lower than threshold) and 25 CAG repeats (at the threshold) [39]. The rate of duplex reannealing for the 25 units was one to six fold slower than the 10 units CAG repeat hairpin, although they were of similar stability. The hairpin structure of 25 units re-formed duplexes reannealed roughly 50-fold slower relative to unstructured random sequences, unstructured scrambled CAG nucleotides, and dinucleotide repeating sequences of identical length [39]. Many more constraints occur in vitro, and whether the lifetime of long flaps exceeds the rate of gap filling synthesis in vitro or in vivo remains to be measured. Nonetheless, the kinetic lifetime of the fold-back structure distinguishes a CAG/CTG tract at the threshold from shorter CAG/CTG tracts by the reannealing rate.

The role of RNA–DNA hybridization in determining the threshold

But could RNA determine the DNA threshold for expansion? Reannealing kinetics appears to be relevant for a TNR threshold mechanism that is R-loop dependent [40,41]. RNA–DNA hybrids form at the expanded (n > 200 rpts) but not normal CGG repeat regions (commonly 30 rpts) in the FMR1 gene from human iPSCs that were differentiated in culture for 30–60 days [40]. The majority of the RNA–DNA duplex occurs between 200 and 300 bp on either side of the expanded CGG tract, consistent with the notion that the promoter harboring the transcribed CGG-repeat tract is the binding site for the FMR1 mRNA. Transcription through the GC-rich FMR1 5′UTR region favors R-loop formation, with the nascent (G-rich) RNA forming a stable RNA:DNA hybrid with the template DNA strand (Figure 2a,b), thereby displacing the DNA strand. Recruitment of the TCR machinery at the stalled site may promote nicking and expansion at the site for repair during removal of the RNA–DNA hybrid block (Figure 2c). In the iPSC system, binding of the FMR1 mRNA to the genomic repeat does not occur before day 45, implying that the hybrid forms slowly [40]. Thus, the size of a stable hybrid might determine the length at which an open transcription bubble ‘sensitizes’ the TNR sensitive to damage (Figure 2a) and render it subject to TCR or BER (Figure 2c). Alternatively, the RNA–DNA bubble may be the threshold ‘impediment’ needed for ‘calling in’ fork reversal [18] or strand-switching [19] resolution mechanisms.

Because of patient variability, it is difficult to determine the precise relationship among transcriptional silencing, the size of the RNA–DNA hybrid, or the level of chemically modified bases. Missing from the iPSC experiments are robust measures of the DNA methylation status and
alterations of the CGG tract length that might have occurred during a 30–60 day differentiation period [40]. Extensive methylation in the promoter region at CGG repeats accompanies transcriptional suppression [42]. If the RNA–DNA hybrid triggers methylation and heterochromatin formation, then another attractive model for expansion is the removal of methylated bases and DNA loop formation via BER [43]. Although removal of methylated bases by BER is accomplished by several DNA glycosylases with different specificities, none are known to promote TNR expansion. In fact, expansion is likely to occur in unmethylated state: (1) Rare individuals having full mutations but normal intelligence lack hypermethylation and maintain expression of FMR1 mRNA [44]. (2) Pharmacologic treatment with the DNA methylation inhibitor 5-aza-2′-deoxycytidine (azadC) reactivates transcription and FMRP expression but does not alter the repeat tract [45]. More likely, expansion is a response to the stress of RNA-induced or protein-induced aggregation/toxicity, which enhances oxidative damage in DNA. Removal of the oxidized bases by the BER or TCR pathways results in loop formation and expansion. Indeed, loss of OGG1 [15**], NEILS 1 [46], and XPA [47] reduces expansion in mice. Novel mechanisms for enhancing oxidative damage and toxicity are discussed below.

Whether RNA–DNA hybrids form at TNRs in other non-coding regions (which generate large expansions) is unknown. In coding regions, the expanded CAG/CTG

### Table 1

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Loop size (n = # rpts)</th>
<th>Pathway</th>
<th>Refs</th>
<th>Endonuclease function</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutL</td>
<td>n &lt; 4 rpts</td>
<td>MMR</td>
<td>[21]</td>
<td>Removal of mismatches and small loops</td>
</tr>
<tr>
<td>MRE11</td>
<td>n.d.</td>
<td>DSBR</td>
<td>[22**]</td>
<td>3′-5′ endonuclease activity at DSBs; hairpin cleaving activity</td>
</tr>
<tr>
<td>MRE11/Sae2</td>
<td>n.d.</td>
<td>DSBR</td>
<td>[22**]</td>
<td>Hairpin cleaving activity with a single strand nick</td>
</tr>
<tr>
<td>APE</td>
<td>n &gt; 4 rpts</td>
<td>BER</td>
<td>[24]</td>
<td>Cleaves 5′ side of abasic site generating a 3′-OH group for polymerase extension</td>
</tr>
<tr>
<td>XPF/ERCC1</td>
<td>n &gt; 4 rpts</td>
<td>NER/TCR</td>
<td>[25,26]</td>
<td>Cleaves a 5′ bubble substrates</td>
</tr>
<tr>
<td>XPG</td>
<td>n &gt; 4 rpts</td>
<td>NER/TCR</td>
<td>[26]</td>
<td>Cleaves a 3′ bubble substrates</td>
</tr>
<tr>
<td>Mus81/Eme1</td>
<td>n.d.</td>
<td>DSBR</td>
<td>[27]</td>
<td>Structure-specific 3′-flap DNA endonuclease that can process substrates resembling replication intermediates</td>
</tr>
<tr>
<td>SLX</td>
<td>n.d.</td>
<td>DSBR</td>
<td>[28,29]</td>
<td>Endonuclease activity toward replication forks, 5′-flaps, and Holliday junctions</td>
</tr>
<tr>
<td>FEN-1</td>
<td></td>
<td>Replication/gap filling repair</td>
<td>[30]</td>
<td>Member of the XPG/RAD2 endonuclease family. Removes 5′ overhanging flaps in DNA repair; processes the 5′ ends of Okazaki fragments in lagging strand DNA synthesis; cofactor in gap filling repair-dependent replication DNA secondary structure can inhibit FEN-1 flap processing</td>
</tr>
<tr>
<td>Gen1/YEN1</td>
<td>n.d.</td>
<td>DSBR, HR</td>
<td>[31]</td>
<td>Holliday junction 5′ flap endonuclease and resolves Holliday junctions</td>
</tr>
<tr>
<td>ZRANB3</td>
<td>n.d.</td>
<td>Replication stress</td>
<td>[32,33]</td>
<td>Localizes to DNA replication sites and interacts with the components of the replication machinery. Maintains genome stability at stalled or collapsed replication forks by facilitating fork restart and limiting inappropriate recombination that could occur during template switching events; acts as a structure-specific endonuclease that cleaves the replication fork D-loop intermediate, generating an accessible 3′-OH group in the template of the leading strand. Cleaves branched DNA structures with unusual polarity (G quartet nuclease 1) cleaves within the single-stranded region 5′ of the barrel formed by stacked G quartets. GQN1 does not cleave duplex or single-stranded DNA, Holliday junctions, or G4 RNA</td>
</tr>
<tr>
<td>GQN1</td>
<td>n.d.</td>
<td>Somatic hypermutability</td>
<td>[34]</td>
<td>(G quartet nuclease 1) cleaves within the single-stranded region 5′ of the barrel formed by stacked G quartets. GQN1 does not cleave duplex or single-stranded DNA, Holliday junctions, or G4 RNA</td>
</tr>
<tr>
<td>SNM1</td>
<td>n.d.</td>
<td>ICL</td>
<td>[35]</td>
<td>Operates predominantly in interstrand crosslink (ICL) repair; structure-specific DNA hairpin opening endonuclease</td>
</tr>
<tr>
<td>RNase H</td>
<td>n.d.</td>
<td>RNA/DNA duplexes</td>
<td>[36]</td>
<td>RNase H is a ribonuclease that cleaves the RNA in a DNA/RNA duplex to produce ssDNA. RNase H is a non-specific endonuclease and catalyzes the cleavage of RNA via a hydrolytic mechanism, aided by an enzyme-bound divalent metal ion. RNase H leaves a 5′-phosphorylated product</td>
</tr>
<tr>
<td>Dicer</td>
<td>RNA</td>
<td>Double stranded RNA</td>
<td>[37**]</td>
<td>Dicer is an endoribonuclease in the RNase III family that cleaves double-stranded RNA (dsRNA)</td>
</tr>
<tr>
<td>Cas9 breaks</td>
<td>RNA</td>
<td>DNA/RNA DNA editing</td>
<td>[38]</td>
<td>RNA-guided DNA endonuclease to generate double-strand</td>
</tr>
</tbody>
</table>
repeat tracts \((n > 35 \text{ rpts})\) overlap in length with those of the FMR-1 ‘normal’ CGG range \([1,2^{*},5^{*},4^{*},5^{*},6^{*}]\) (commonly 30 rpts), which does not form hybrids. Moreover, CAG expansions do not impose transcription silencing of their respective genes \([1,3]\). If a minimum DNA–RNA hybrid causes the transcriptional silencing at a threshold length, then it is unlikely to be a mechanism that is common to all TNR genes.

**Double stranded RNA models for expansion.**

Another consideration in a RNA-dependent hybridization model for threshold is the effect, if any, of bi-directional transcription of the TNR region \([48^{*}]\). For example, several novel anti-sense FRM1 transcripts exist in the FRM1 locus (ASFMR4-6), and some overlap the CAG repeat region \([49]\). ASFMR4 transcript is spliced, polyadenylated and exported to the cytoplasm \([42,49]\). If a bi-directional transcript overlies with the sense transcript, double stranded RNA is formed as a Dicer substrate. It is not easy to imagine how short siRNA hybrids within the TNR tract results directly in expansion. Either multiple siRNA binding creates a RNA–DNA hybrid of similar length to that of an mRNA hybrids \([46]\), and are removed by similar mechanisms, or the shorter RNA–DNA hybrid opens the DNA sufficiently to increase exposure to oxidative DNA damage at a preferred threshold length (Figure 2a).

**Protein–RNA interaction models for the threshold**

New models provide insight on how RNA–protein complexes of threshold length might provoke chemical lesions in DNA, and lead to expansion. TAR-DNA-binding protein 43 (TDP-43) \([50]\) is poised to bind to a RNA–DNA hybrid. TDP-43 is a dimeric protein with two RNA recognition motif (RRM) domains that bind both DNA and RNA \([50,51^{*},52]\) (Figure 3a–c), and interact with fragile X mental retardation protein (FMRP) in an (FMRP)/Staurop (STAU1) complex \([53]\). This complex forms aggregates analogous to those of polyglutamine
proteins, which induce cellular stress and oxidative DNA damage. The DNA length at which the encoded RNA forms aberrant protein–RNA complexes may be the threshold for the enhanced stress.

The mechanisms of RNA aggregate formation are unknown, but it is likely due to the disruption of complex formation at its C-terminus. TDP-43 interacts at its C-terminus with the hnRNP family of translation factors, as well as the splicing factors muscleblind (MBNL) and CUG-BP1 (CUG binding protein 1) [54]. MBNL and CUG-BP1 impart two opposing effects on splicing, and they occur through binding of distinct regions of the target RNA [55]. Both CUG-BP1 and MBNL bind to short-structured CUG and CCUG repeats in RNA with high affinity and specificity [55] (Figure 3d,e). Only 6 base pairs are necessary for MBNL binding; two pyrimidine mismatches and four guanosine–cytosine base pairs that form in a helical region of a stem-loop in the endogenous pre-mRNA target [55] (Figure 3e). In the myotonic dystrophy gene (DM1), these two regions of the RNA reside on the 3′ and 5′ sides that surround the TNR [56].

The length of the TNR tract affects only MBNL binding and impairs its function. A loss-of-function in MBNL and a gain-of-function in CELF4 tend to favor generation of the alternatively spliced forms.

TDP-43 also binds to both the 3′ and 5′ end of the DM1 mRNA, and raises the possibility of that binding of MBNL and TDP-43 occurs at the same sites. Whether these two proteins overlap in the recognition to mRNA is unknown, but the common binding sites and functionality in the DM1 mRNA raise the possibility that the bi-partite mRNA binding at the C-terminus of TDP-43 integrates translation and splicing activity. Interestingly, TDP-43 controls its own expression through a negative feedback loop involving interactions with its mRNA at the 3′ end [57]. Furthermore, the domain structure of TDP-43 is similar to that of both heterogeneous nuclear ribonucleoprotein (hnRNP) and muscleblind (MBNL)
[58] (Figure 3f): an N-terminal domain (NTD) and two tandem RNA recognition motifs (RRM1 and RRM2), followed by a C-terminal glycine-rich region (G) (Figure 3a–c). The C-terminus of TDP-43 acts as a hub that regulates both splicing and translation. Indeed, TNR coding transcripts are associated with an unusual type of translation, Repeat Associated Non-ATG translation (RAN-translation) [59**]. RAN-translation does not require an ATG translation start site, and random translation at TNRs occurs in all reading frames [59**].

Given its hub-like features, maintaining the C-terminus of TDP-43 would appear to be a key regulatory process. Indeed, pathological TDP-43 in the cytoplasmic and intranuclear inclusions is hyper-phosphorylated, ubiquitinated, and cleaved to ~25 kDa C-terminal fragments in affected brain regions [60]. C-terminal-deleted TDP-43 without the glycine-rich tail is sufficient to form a head-to-head homodimer primarily via its N-terminal domain, which form fibrils in vitro [60]. Thus, proteolytic cleavage of TDP-43 within the RRM2 removes the N-terminal dimerization domain and produces unassembled truncated RRM2 fragments, which can abnormally oligomerize into high-order inclusions (Figure 3). The resulting increase in oxidative DNA damage promotes expansion indirectly by RNA-mediated depletion of TDP-43/FMRP/STAU1 in the nucleus and an increase in cellular stress.

The role of maintaining balanced DNA methylation in expansion

Whether this type of RNA-mediated mechanism applies to all triplet repeat disorders is unknown, but there are direct links between them and mitochondrial metabolism. A TDP-43 binds to the mRNA of the silent information regulator 1 (SIRT1), which is implicated in double-stranded DNA break repair and DNA metabolism in all cells [61]. SIRT1 is an NAD⁺ dependent class III histone deacetylase [61], which cooperates with elongation factor 1 (E2F1) to regulate apoptotic response to DNA damage. SIRT1 knockdown results in poly Q-expanded aggregation of androgen receptor (AR) and α-synuclein [62], consistent with a role of the SIRT1mRNA-TDP-43 complex in aggregation, and supports the notion that RNA processing by TDP-43 and chromatin organization SIRT1 are functionally connected. TDP-43 regulates alternate splicing of the CFTR...
RNA at the intron8/exon9 junction, implying that alternative splicing may have a direct consequence on the chromatin organization, which is altered at long, congenital TNR lengths.

Interestingly, isocitrate dehydrogenase 1 (IDH1) and IDH2 catalyze the interconversion of isocitrate and α-ketoglutarate (α-KG) [63] (Figure 4a). α-KG is a TCA cycle intermediate in mitochondria, and is an essential co-factor for many enzymes, including JmjC domain-containing histone demethylases [63,64**], and a family of 5-methylcytosine (5mC) hydroxylases, Ten-eleven translocation dioxygenase (TET) [64**] and EglN prolyl-4-hydroxylases (Figure 4a). Both TET1 and TET3 proteins contain a DNA-binding motif that is believed to target CpG sites (Figure 4a). TET2 converts 5-methylcytosine to 5-hydroxymethylcytosine (5-hmC) in DNA and uses α-ketoglutarate as a co-substrate [65]. The resulting (5-hmC) is removed by the BER enzyme thymine DNA glycosylase (TDG) [64**] (Figure 4b). At the excision site, cytosine replaces 5-hmC, and methylation occurs subsequently to restore the methylated state and 5-mC [64**] (Figures 4b and 5a,b). Thus, metabolism is apparently a regulatory mechanism to maintain a balanced methylation state, and influences expansion. Since methylation status does not appear to play a role in expansion per se, RNA-induced and protein-induced toxicity may act in a feed-back loop, producing a toxic oxidation cycle and expansion during removal of the oxidative DNA damage (Figure 5c).

**Conclusions**

Although new possibilities for DNA-mediated, RNA-mediated and protein-mediated toxicity are emerging, these diverse pathways, in the end, are likely to induce expansion by similar mechanisms (Figure 5). Physically, expansion occurs by loop formation at free DNA ends during DNA excision, by polymerase slippage or by

---

Two-state threshold model. (a) The binding pocket of the OH-CH3-cytosine in TDG. (b) The overall cartoon structure of the OH-CH3-cytosine-TDG complex. (c) The toxic oxidation cycle. We propose a two-state model for expansion. Expansion arises from toxicity imparted from RNA and protein-mediated toxicity by two mechanisms. RNA and protein-mediated toxicity induces mitochondrial stress and a concomitant increase in oxidative damage to DNA. The oxidative damage is removed by DNA glycosylases. 8-oxo-guanine glycosylase (OGG1) is a major enzyme that removes oxidative damage, but it can also be removed by the NEIL1 1’ glycosylase, or the machinery of TCR'. Single strand break intermediates arise during base removal and produce flaps that fold-back to generate structural intermediates for expansion.
strand switching events that occur during replication or fork-reversal. From this simple viewpoint, we can construct both physical and functional definitions of an expansion threshold. Physically, the threshold defines a kinetic point in which self-pairing ‘wins’ over duplex reformation. Structures form at Okazaki fragment ends and/or at single strand breaks are trapped by gap filling synthesis or continued replication (Figure 5). Functionally, the threshold is likely to be the limiting length at which lesion load induces DNA repair.

Acknowledgement
This work was supported by National Institutes of Health grants ES020766-01 (to CTM), and CA092584 (CTM), and NS060115 (CTM).

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


Excellent review on TCR.


First report linking oxidative DNA damage directly to a mechanism expansion via a base excision repair.


Novel model for a small chemical lesion, 8-oxo-G, to lesion by-pass.


First report demonstrating the mechanism of endonuclease switching of MRE11/Rad50.


Trinucleotide expansion in disease Lee and McMurray 139


Excellent summary of the discovery and important cooperative activities of muscleblind and CUG bindin protein.


This report reveals a clear relationship between TDP-43 and its binding to RNA.


Discovery of novel form of translation in TNR repeats.


Excellent review of methylation cycle in DNA.