Can MSH3 lowering stop HTT repeat expansion in its CAG tract?

Ross Ferguson and Sarah J Tabrizi

Huntington's Disease Centre, Department of Neurodegenerative disease, UCL Queen Square Institute of Neurology, University College London, WC1N 3BG, UK; Dementia Research Institute at UCL, London WC1N 3BG, UK.

Corresponding author: s.tabrizi@ucl.ac.uk

Huntington's disease (HD) is a neurodegenerative repeat expansion disorder caused by an expansion of the CAG trinucleotide repeat present in exon 1 of the huntingtin (HTT) gene. The expanded repeat exhibits instability in somatic tissues resulting in an inexorable increase in length over the HD gene carrier's lifetime. Although the length of the expanded CAG repeat is the strongest predictor of age at onset and progression, potential genetic modifiers have also been identified[1,2]. Many of these modifiers are within DNA repair genes and have been implicated in repeat expansion, including the mismatch repair protein MSH3. The process of somatic repeat expansion in the brain appears to be a major factor in driving disease progression leading to production of more toxic HTT protein species. As a monogenic disease, many genetic therapeutic approaches past and present aim to ameliorate the effects of the mutant HTT protein by lowering its levels. Here, O'Reilly and colleagues[3] demonstrate the potential of an alternative therapeutic intervention to slow the expansion of the CAG repeat by targeting MSH3. They achieve suppression of somatic expansion in two different but complementary HD mouse models by instead lowering the expression MSH3, using a nucleic acid therapeutic modality established previously by the Khvorova group, the di-valent siRNA (di-siRNA)[4].

The successful slowing of expansion by targeting MSH3 is an elegant example of utilising human genetic variation to lead in identifying targets tolerant of modulation. Polymorphisms in *MSH3* that are predicted to reduce function were associated with slower disease progression in HD cohorts[1].

These polymorphisms are 9bp tandem repeat sequences of varying length in exon 1. The three repeat variant (termed 3a) showed reduced expression and was associated with reduced somatic expansion, delayed onset and slower progression[5]. While MSH3 has long been known to play a role in repeat expansion using mouse and *in vitro* models, these naturally occurring human variants suggested that reduced MSH3 function is tolerated at a level sufficient to impact upon repeat expansion, making it an attractive therapeutic target.

The canonical function of MSH3 in mismatch repair is to maintain genomic stability through detection and repair of the short stretches of mismatched DNA that occur routinely during transcription and replication. MSH3 also plays a role in instability at the HTT CAG repeat tract, a process which may ultimately lead to, or exacerbate, HD associated neuronal dysfunction. Any therapeutic which slows this process in HD could also be relevant to many other repeat expansion disorders with unstable repeats, such as the various spinocerebellar ataxias or fragile X syndrome.

Proficient MMR in mammalian cells is dependent on the MutS protein homologs MSH2, MSH6 as well as MSH3. These proteins form heterodimer complexes of either MSH2 & MSH6 (MutSα) or MSH2 & MSH3 (MutSβ). While the MutSα complex preferentially drives repair of single base mismatches and 1-4bp loop-outs, MutSβ (and so MSH3) is responsible for recognising longer extrusions[6]. Mismatch repair by either of the MutS complexes starts by recruiting MutL protein homologs, another family of heterodimeric complexes, to form a ternary complex (Figure 1A). MLH1 forms the common component of three different MutL heterodimers with either PMS1, PMS2 or MLH3, each with differing but overlapping roles in MMR. MutL nuclease activity nicks the DNA flanking the mismatch which is then removed by an exonuclease such as EXO1, before resynthesis and ligation by DNA polymerase δ and LIG1.

In the context of somatic instability, *in vitro* studies suggest MutSβ associates with extra-helical DNA structures. These structures occur with higher frequency in highly repetitive sequences such as the HTT CAG repeat due to strand slippage, polymerase stalling and R-loop formation during

transcription or replication [7]. The subsequent aberrant repair attempts by the MMR factors recruited by MutSβ lead to the inclusion of additional repeats in the strand opposing the looped-out DNA upon its resolution (Figure 1B). While the mechanism and the influence of other DDR factors is not completely characterised (reviewed by Miller and Usdin[6]), it is clear from multiple models that lowering MSH3 and therefore MutSβ slows this process[8].

Distribution and uptake of nucleic acid therapeutics is an ongoing challenge. While the composition of antisense oligonucleotides (ASOs) can be heavily modified to improve biodistribution, similar modification of siRNAs can adversely affect their loading into the RISC complex and efficacy. This is exacerbated upon conjugation to ligands such as cholesterol to improve cellular uptake. Di-siRNAs overcome this by utilising two covalently bound siRNAs with partially phosphorothionated (PS) backbones[4]. The relative PS content of these di-siRNAs is increased compared the maximum possible in a singular siRNA, allowing for superior distribution within the brain without compromising efficacy.

The authors designed a panel of di-siRNAs targeting MSH3 using their previously validated chemistry. The efficacy of these 60 compounds were tested first in mouse, human and non-human primate cell lines to assess cross-species target engagement. As the effects of knocking out MSH3 on CAG repeat somatic expansion has been investigated previously in HD mouse models, they selected di-siRNA with good efficacy in both mouse and human to test whether they could lower MSH3 expression sufficiently to attenuate somatic expansion in the same models.

The Hdh^{Q111} HD mouse line exhibits somatic instability and was generated by the MacDonald group by knocking a 109 CAG repeat into the mouse HD homolog (*Hdh*) gene. Instability is suppressed in Hdh^{Q111} mice crossed to an *Msh3* knock-out line[9]. It is also suppressed in the absence of *Msh2*, but not *Msh6, s*upporting the dominant role of the MSH3-containing MutSβ complex in expansion. The authors introduced their lead di-siRNAs by intracerebroventricular injection to 12-week-old mice. Treatment reduced MSH3 expression in the striatum, cortex & thalamus of Hdh^{Q111} mice along with a reduction in the instability of the repeat in striatal tissue. These results were reproduced in the BAC-

CAG HD mouse, another HD mouse model with somatic instability generated by the Yang Lab[10]. This mouse carries a full length human HTT allele with 120 uninterrupted CAG repeats and flanking regulatory regions. Lowering MSH3 expression results in the reduced somatic instability in every experiment.

CAG repeat expansion is slowed in both *Hdh*^{Q111} and BAC-CAG models with di-siRNA treatment however the endpoint levels of MSH3 protein and transcript suggest a waning of silencing over the 20-28 week interval post-injection. Despite this, the period of efficacy post-injection was clearly sufficient to suppress a substantial degree of CAG repeat expansion in vulnerable tissues. Unlike a genetic knock-out, the di-siRNAs do not completely ablate MSH3 yet they still appear to have a dramatic effect on instability, a promising indicator for future pharmacodynamic considerations in translation. Our knowledge of MSH3 effects in somatic instability primarily comes from mouse models and cell lines. As such, it will be of great interest to see this effect recapitulated in human *ex vivo* models of somatic expansion, such as HD iPSC derived neurons and specifically striatal neurons carrying expanded repeats in the endogenous HTT gene.

One concern for this approach is that by targeting a DNA damage pathway we are potentially increasing overall genomic instability and cancer risk. Heterozygous loss of function mutations in many MMR proteins are associated with hereditary colorectal cancers. These include MSH2 and MSH6, as well as GWAS hits such as MLH1/PMS2 but importantly not MSH3[11]. O'Reilly did not observe increased microsatellite instability during their experiments here, nor have increased incidences of cancer or reduced longevity been reported for MSH3 knock-out mice[12]. In humans, homozygous loss of MSH3 has been associated with colorectal cancer[11] but it has not been conclusively linked to cancers of the brain. As such non-systemic, partial lowering of MSH3 in the brain is likely a relatively safe approach, further evidenced by the prevalence of functional variants in the general population, as highlighted by Moss and colleagues[1] from published Exome Aggregation Consortium data. Similar considerations could be applied to other HD risk modifiers which also tolerate loss of function. These data suggest a

promising therapeutic approach with which to target somatic CAG repeat expansion in humans, and with hope, slow disease progression.

Declaration of Interests

In the past year, through the offices of UCL Consultants Ltd, a wholly owned subsidiary of University College London, SJT has undertaken consultancy services for Alnylam Pharmaceuticals, Annexon, Ascidian Therapeutics, Arrowhead Pharmaceuticals, Atalanta Therapeutics, Design Therapeutics, F. Hoffman-La Roche, HCD Economics, IQVIA, Iris Medicine, Latus Bio, LifeEdit, Novartis Pharma, Pfizer, Prilenia Neurotherapeutics, PTC Therapeutics, Rgenta Therapeutics, Takeda Pharmaceuticals, UniQure Biopharma, Vertex Pharmaceuticals. In the past 12 months, University College London Hospitals NHS Foundation Trust, SJTs host clinical institution, received funding to run clinical trials for F. Hoffman-La Roche, Novartis Pharma, PTC Therapeutics, and UniQure Biopharma.

References

1. Moss, D. J. H. *et al.* Identification of genetic variants associated with Huntington's disease progression: a genome-wide association study. *Lancet Neurol.* **16**, 701–711 (2017).
2. Lee, J.-M. *et al.* CAG Repeat Not Polyglutamine Length Determines Timing of Huntington's Disease Onset. *Cell* **178**, 887-900.e14 (2019).
3. O'Reilly. Di-valent siRNA Mediated Silencing of MSH3 Blocks Somatic Repeat Expansion in Mouse Models of Huntington's Disease.
4. Alterman, J. F. *et al.* A divalent siRNA chemical scaffold for potent and sustained modulation of gene expression throughout the central nervous system. *Nat. Biotechnol.* **37**, 884–894 (2019).
5. Flower, M. *et al.* MSH3 modifies somatic instability and disease severity in Huntington's and myotonic dystrophy type 1. *Brain* **142**, 1876–1886 (2019).
6. Miller, C. J. & Usdin, K. Mismatch repair is a double-edged sword in the battle against microsatellite instability. *Expert Rev. Mol. Med.* **24**, e32 (2022).
7. Xu, P., Pan, F., Roland, C., Sagui, C. & Weninger, K. Dynamics of strand slippage in DNA hairpins formed by CAG repeats: roles of sequence parity and trinucleotide interrupts. *Nucleic Acids Res.* **48**, 2232–2245 (2020).
8. Keogh, N., Chan, K. Y., Li, G.-M. & Lahue, R. S. MutSβ abundance and Msh3 ATP hydrolysis activity are important drivers of CTG•CAG repeat expansions. *Nucleic Acids Res.* **45**, 10068–10078 (2017).
9. Dragileva, E. *et al.* Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol. Dis.* **33**, 37–47 (2009).
10. Gu, X. *et al.* Uninterrupted CAG repeat drives striatum-selective transcriptionopathy and nuclear pathogenesis in human Huntingtin BAC mice. *Neuron* **110**, 1173-1192.e7 (2022).
11. Adam, R. *et al.* Exome Sequencing Identifies Biallelic MSH3 Germline Mutations as a Recessive Subtype of Colorectal Adenomatous Polyposis. *Am. J. Hum. Genet.* **99**, 337–351 (2016).
12. de Wind, N. *et al.* HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. *Nat. Genet.* **23**, 359–362 (1999).

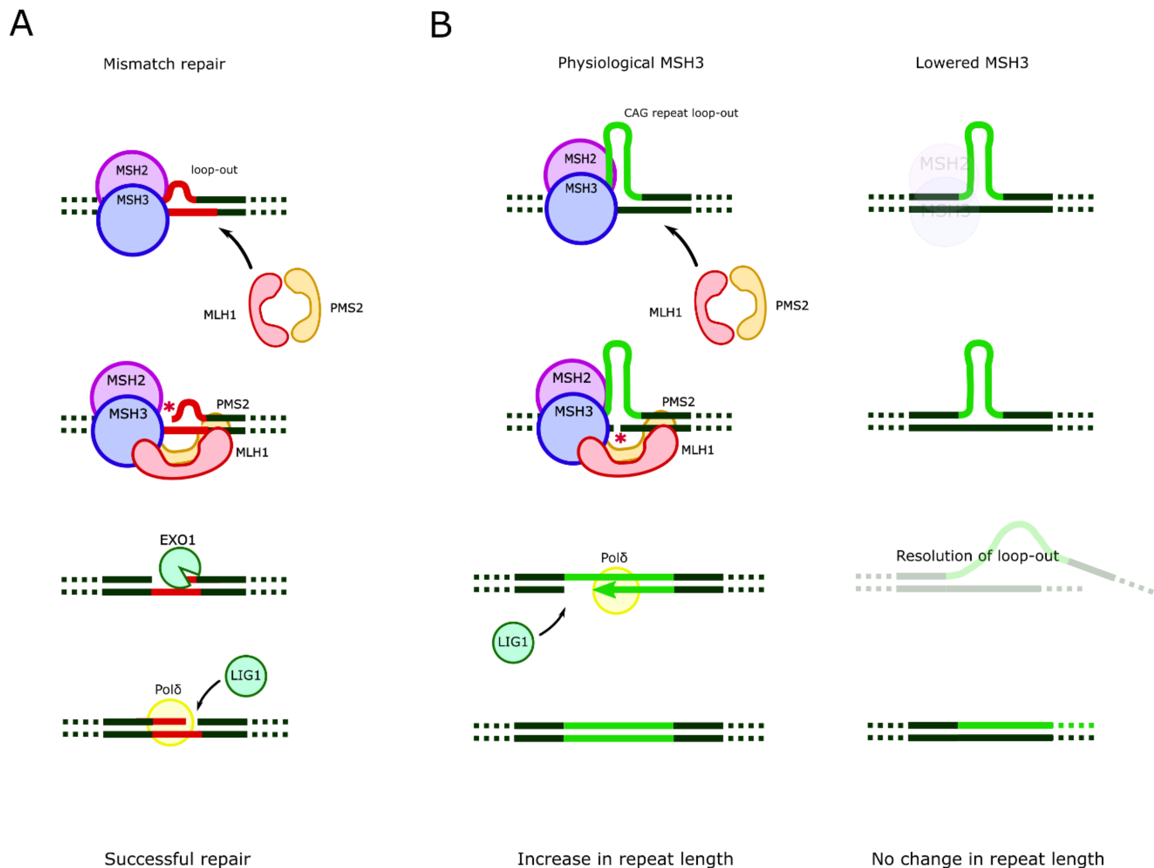

Figure 1. Mismatch repair and its role in repeat expansions. Mismatch repair is initiated by detection of mispairing by two heterodimeric MutS complexes. MutSα (MSH2 & MSH6) recognises single base mismatches and 1-4bp loop-outs, while the MutSβ complex (MSH2 & MSH3) detects small 2-10bp DNA extrusions (A). Detection results in the recruitment of MutL family heterodimers to form a ternary complex. The endonuclease activity of MutL factor PMS2 nicks the DNA in the strand harbouring the incorrect sequence. This allows loading of Exo1 which excises nucleotides 5' to 3' until it reaches a flanking nick. The resulting gap is then filled by DNA polymerase delta and nicks ligated closed by LIG1. B) CAG loop-out structures form at the expanded repeat in exon 1 of the huntingtin gene during replication and transcription. These loop-outs are detected by MutSβ (MSH2 & MSH3) which then recruits MutL heterodimers such as MutLα (MLH1 & PMS2). The strand opposing the loop-out is nicked (asterisk) and resected. The additional bases comprising the loop-out are incorporated into the opposing strand during synthesis and ligation, increasing the length of the repeat. Where MSH3 is absent due to knock-out, or reduced in levels due to divalent-siRNAs targeting it as O'Reilly et al., have shown in their manuscript, this process cannot occur. The loop-outs are resolved through other means maintaining the pre-existing length of the CAG repeat.