TREX1 DNA exonuclease deficiency, accumulation of single stranded DNA and complex human genetic disorders

O'Driscoll, M (2008) TREX1 DNA exonuclease deficiency, accumulation of single stranded DNA and complex human genetic disorders. DNA Repair, 7 (6). pp. 997-1003. ISSN 1568-7864

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/1810/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
‘Hot Topic’ Review.

TREX1 DNA exonuclease deficiency, accumulation of single stranded DNA and complex human genetic disorders.

Mark O’Driscoll

Genome Damage & Stability Centre
University of Sussex
Falmer
Brighton
East Sussex
BN1 9RQ
UK

Ph. 0044 (0) 1273 877 515
Fax. 0044 (0) 1273 678 121
m.o-driscoll@sussex.ac.uk

Running title: TREX1, Aicardi-Goutieres syndrome, single stranded DNA, DNA damage response.
**Abstract.**

Aicardi-Goutieres syndrome (AGS) is an unusual condition that clinically mimics a congenital viral infection. Several genes have recently been implicated in the aetiology of this disorder. One of these genes encodes the DNA exonuclease TREX1. Recent work from Yang, Lindahl and Barnes has provided insight into the cellular consequence of TREX1-deficiency. They found that TREX1-deficiency resulted in the intra-cellular accumulation of single stranded DNA resulting in chronic activation of the DNA damage response network, even in cells from Trex1-mutated AGS patients. Here, I summarise their findings and discuss them in context with the other AGS causative genes which encode subunits of the RNase H2 complex. I describe mechanisms by which the inappropriate intra-cellular accumulation of nucleic acid species might deleteriously impact upon normal cell cycle progression. Finally, using the example of Systemic Lupus Erythematosus (SLE), I also summarise the evidence suggesting that the failure to process intermediates of nucleic acid metabolism can result in the activation of uncontrolled autoimmunity.
1. Introduction.

In 1984 two French paediatric neurologists, Jean Aicardi and Francoise Goutieres, described an unusual condition in a Portuguese family characterised by chronic lymphocyte infiltration into the cerebrospinal fluid (CSF), calcification of the basal ganglia and early onset progressive severe encephalopathy (brain degeneration) [1]. Since the clinical sequelae of this condition closely mimicked intra-uterine infection, cytomegalovirus was originally suspected as the causative agent. But, detailed serological screening for TORCH (toxoplasmosis, rubella, cytomegalovirus and herpes simplex) failed to implicate any infectious agent. What made this condition even more intriguing was the birth of a second affected child to the consanguineous parents, strongly suggesting an autosomal recessive disorder. Subsequent studies on other patients confirmed the heritable nature of this condition and extended the clinical spectrum to include progressive microcephaly (reduced brain size), cutaneous lesions on the extremities (fingers, toes, ears) reminiscent of chilblain’s and increased levels of α-interferon (α-INF) in the CSF and blood [2]. Since its original description in 1984, several Aicardi-Goutieres syndrome (AGS) susceptibility loci have been identified from mapping studies in different families (Table 1). In the Summer of 2006, following extensive genetic analysis of multiple families, causative mutations were described for AGS in several distinct genes, one of which encoded the major 3’-5’ DNA exonuclease in mammalian cells DNase III/TREX1 (Three prime repair exonuclease I) (Table 1)[3] [4]. Recent work from the Yang, Lindahl and Barnes has uncovered an exciting link between defective DNase III/TREX1 (hereafter referred to TREX1), the accumulation of excessive single stranded DNA (ssDNA) in the cytoplasm and chronic activation of the ATM-dependent DNA damage response [5]. This work follows-on directly from
observations from the Lieberman group showing that DNase III/TREX1 resides within a complex with NM23-H1 endonuclease where they function together to degrade DNA in a caspase-independent cell death pathway [6]. Together these findings provide information on the biologically relevant function of TREX1 and hint at a provocative mechanistic explanation for how mutations in a DNA exonuclease can result in a complex human disorder that mimics a congenital viral infection.

2. A brief summary of mammalian DNA exonucleases.

The majority of DNA exonuclease activity in mammalian cell extracts derives from two distinct nuclear enzymes originally designated DNase III (3’-5’) and DNase IV (5’-3’). Exonuclease I (EXO1) also encodes a 5’-3’ exonuclease that functions mainly in mismatch repair and recombination [7]. These activities are distinct from the DNA endonucleases DNase I and DNase II. DNase IV was subsequently renamed Flap Endonuclease 1 (FEN1), and was shown to function in the processing of 5’ single stranded overhangs that arise during lagging-strand DNA replication, DNA repair and recombination [8,9]. DNase III (TREX1) was found to encode a non-processive 3’-5’ DNA exonuclease with a preference for ssDNA or mispaired 3’ termini [10,11]. TREX1 shares homology to the *E. coli* DnaQ/MutD editing subunit of the replicative DNA polymerase III holoenzyme [12]. Since two of the principal mammalian nuclear DNA polymerases, Polα and Polβ, do not possess intrinsic proofreading 3’exonuclease activity, it was originally thought that TREX1 may serve to excise mismatched dNTP’s during lagging strand DNA synthesis (Polα-dependent) or gap filling during base excision repair (BER, Polβ-dependent) [13,14]. In fact, TREX1 was shown to remove a 3’mismatch in an *in vitro* BER assay reconstituted with Polβ [12]. The functional/biological significance of this remained unclear as
other 3’ DNA exonuclease activities, including one attributed to APE1/HAP1 apurinic-apyrimidinic (AP) endonuclease, have been implicated in BER.

TREX2 shares significant (approx 44%) amino acid identity with TREX1, although it lacks the C-terminal 68 amino acid extension found in TREX1. Similarly to TREX1, TREX2 also functions as a 3’-5’ exonuclease [13]. The precise biological function of TREX2 is currently unclear although a possible role in mediating genomic stability has been suggested [15]. The crystal structure of Trex1 has recently been described and readers are referred to these studies for in-depth description of TREX1 structure-function analysis [16,17].

3. A mouse model for TREX1-deficiency.

A direct, biologically relevant role for TREX1 in BER was further put in doubt when TREX1-/- mice generated in the Big Blue indicator background failed to exhibit an increased spontaneous mutation frequency [18]. In fact, TREX1-/- mice did not exhibit increased tumour incidence. Furthermore, TREX1-/- mouse embryonic fibroblasts (MEFs) did not exhibit an increased sister chromatid exchange frequency (typical of XRCC1-defective hamster cells) or sensitivity towards hydrogen peroxide or ionising radiation (IR) [18]. Although TREX1-/- mice were born at Mendelian frequencies, suggesting that the gene was non-essential, they exhibited a significantly reduced life span with spontaneous development of inflammatory myocarditis (inflammation of the muscle wall of the heart), progressive dilated cardiomyopathy and consequent circulatory failure [18]. The inflammatory myocarditis found in TREX1-/- animals was consistent with an autoimmune aetiology since it occurred in pathogen-free animals.
4. TREX1-deficiency, ssDNA and the DNA damage response.

The exact mechanism of TREX1-deficient disease has remained obscure despite causing AGS in human [4]. Using TREX1-/- MEFs and AGS-derived patient cells, Yang et al have uncovered interesting cellular consequences of TREX1-deficiency that help shed light on this issue [5]. Firstly, they corroborated the recent observation by the Lieberman group that TREX1 predominantly localises to the cytoplasm, specifically the endoplasmic reticulum (ER), which is somewhat unexpected for a DNA-specific exonuclease [5,6]. Interestingly, TREX1 was found to localise to the nucleus at replication (BrdU positive) foci following DNA damage. They also found that TREX1-/- MEFs do not progress normally through the cell cycle, instead exhibiting a reduced S-phase population and accumulation in G1.

Furthermore, following treatment with IR, TREX1-/- MEFs failed to arrest properly in S and G2/M-phases of the cell cycle. Molecular dissection of these cellular phenotypes showed that TREX1-/- MEFs exhibited spontaneously increased levels of both p53 and p21 suggestive of chronic activation of the DNA damage response machinery. Interestingly, TREX1-/- MEFs also exhibited a significantly reduced endogenous level of CHK2. Chemical inhibition of ATM kinase activity using the specific ATM inhibitor, KU-55933, restored CHK2 expression levels. Although the mechanism remains unclear, collectively these data indicate chronic ATM-mediated checkpoint activation in TREX1-/- MEFs in the absence of exogenous DNA damage [5].

Since TREX1 was found to localise to BrdU positive foci in S phase cells exposed to DNA damage, coupled with its known preference for ssDNA as a substrate, Yang et al reasoned that TREX1 likely acts on ssDNA generated during
DNA replication and that this DNA species may accumulate in TREX1-/- cells [5]. Using an antibody specific to ssDNA, TREX1-/- MEFs showed increased levels of ssDNA but, strikingly, it was almost exclusively extranuclear. Co-localisation with the ER-specific marker calreticulin showed that the ssDNA was associated with the ER. Furthermore, cellular fractionation studies showed that cytoplasmic extracts from TREX1-/- MEFs specifically contain ssDNA polynucleotides of between 60-65 bases in size. Using BrdU labelling under non-denaturing conditions, Yang et al proposed that the ER-associated ssDNA was derived from S phase cells [5]. Moreover, primary fibroblasts from AGS patients with a homozygous mutation in Trex1 exhibited a similar cell cycle profile as TREX1-/- MEFs (low S-phase and accumulation at G1), along with increased p53 and p21 levels, reduced CHK2 expression and increased ER-associated ssDNA. Together, these data suggest that TREX1-deficiency results in the accumulation of excessive amounts of short ssDNA oligonucleoide within the cytoplasm of replicating cells, which somehow results in chronic activation of the DNA damage response-signalling network [5].

5. Is there a functional link between TREX1-deficiency and the other AGS causative genes?

AGS mimics congenital viral infection. In fact, exclusion of perinatal infection is one of the diagnostic criteria for this disorder [1,2]. TREX1-deficiency leads to the inappropriate accumulation of ssDNA within the cell, a factor presumed to contribute to the aetiology and progression of this condition (see below). Why should or how can excessive amounts of DNA within a cell result in a pathological condition? Furthermore, what of the other AGS causative genes (Table 1)? Do they also play a role in processing nucleic acids? Crow and colleagues, along with implicating
TREX1-deficiency in AGS, also identified mutations in each of the mammalian subunits of the RNase H2 complex in several non-Trex1-mutated AGS families [3,4]. The equivalent complex in *S. cerevisiae* (Rnh2Ap- Rnh2Bp- Rnh2Cp) endonucleolytically cleaves ribonucleotides from RNA:DNA duplexes [19]. The RNase H2 complex is thought to specifically function in the removal of lagging-strand Okazaki fragment RNA primers during DNA replication as well as in the excision of mis-incorporated ribonucleotides from DNA duplexes [19]. Therefore, the identification of further AGS causative mutations in genes that co-ordinately function as a nuclease (ribonuclease in this case) reinforces the notion that the potential excessive intra-cellular accumulation of nucleic acid species (ssDNA, DNA:RNA duplexes, even ribonucleotide containing DNA:DNA duplexes) may contribute to the pathology of AGS [3-5].

6. TREX1 and cell death: cutting DNA for a reason.

TREX1 activity has also been implicated in a specific caspase-independent cell death pathway utilised by cytotoxic T cells and Natural Killer cells during the immune response [6]. Granzyme A (GzmA) is a serine protease that once delivered into target cells by perforin via immunological synapse initiates a cell death pathway that morphologically resembles caspase-dependent apoptosis [20,21]. GzmA mediated cell death is characterised by the appearance of ssDNA, as opposed to the dsDNA fragmentation typical of apoptosis. Once GzmA is introduced into cells it cleaves SET, an inhibitor of the NM23-H1 endonucelase resulting in NM23-H1 activation [22]. SET and NM23-H1 reside in the ER in a large complex (SET complex) one of whose components has recently been shown to be TREX1 [6]. It is thought that NM23-H1 and TREX1 work co-ordinately whereby activated NM23-H1
introduces single stranded nicks into DNA and TREX1 subsequently removes nucleotides from the nicked 3’ end to prevent the BER pathway from rapidly rejoining the nicks [6]. The SET complex is thought to play an important role in the defence against viruses and tumour cells that have evolved strategies for evading caspase-mediated apoptosis [23]. Silencing of *Trex1* (siRNA) can inhibit GzmA-mediated cell death [6]. Therefore, TREX1 clearly has a role in some cell types outside of processing DNA replication intermediates.

7. **A mechanism for a role of TREX1-deficiency and RNase H2-deficiency in autoimmune disease.**

A member of the Toll-like receptor family, TLR9, is associated with the ER of several immune system cell types (eg. macrophages), where it responds to endocytosed DNA (including ssDNA species) containing unmethylated CpG residues [24,25]. TLR9 normally serves as a pathogen sensor by detecting unmethylated DNA that is characteristically over-represented in bacterial DNA compared to mammals [25]. The S-phase-derived ER-associated ssDNA observed by Yang *et al* to accumulate in TREX1-/- MEFs and AGS patient cells is unlikely to have undergone extensive de novo methylation at CpG residues [5]. Therefore, it is possible that it may result in, or contribute to uncontrolled activation of the innate immune system with pathological consequences (autoimmune reaction). Similarly, as Trex1-deficiency inhibits GzmA-mediated cell death, this could indirectly result in the failure to remove autoreactive cytotoxic T cells whose accumulation could also contribute to autoimmunity [6].
A diagnostic clinical feature of AGS is elevated levels of $\alpha$-IFN in the CSF and blood suggestive of viral infection [2]. It is known that dsRNA and dsDNA can activate the innate immune system and induce $\alpha$-IFN production [26]. Reverse transcription, involving the formation of DNA:RNA hybrids, is an essential step during retroviral propagation. Failure of the RNase H2 complex to adequately clear DNA:RNA Okazaki fragments could conceivably result in the accumulation of nucleic acid species that mimic a retroviral infection. Although, it should be noted that the accumulation of intracellular nucleic acids in non-$Trex1$ mutated AGS patient cells has not yet been formally demonstrated.

How does chronic activation of the DNA damage response impact on immune system function? An elegant study from Gasser and colleagues has shown that some stimulatory receptors (NKG2D receptors) of the innate immune system are activated in an ATM, ATR and CHK1-dependent manner [27]. A proposed role for this NKG2D receptor up-regulation is to enable the innate immune system to distinguish between normal (self) and diseased (infected/transformed cells) [28]. Interestingly, chronic activation of the DNA damage response has recently been found to precede malignant transformation [29-31]. It is particularly interesting to note that cancer predisposition has not been observed in AGS despite several patients being in the third decade of life [2]. It is tempting to think that this may be because the innate immune system in AGS individuals is already primed or activated (inappropriately) to target non-self (i.e. transformed) cells.

Systemic Lupus Erythematosus (SLE) is a multifactorial, genetically heterogeneous, systemic autoimmune disease characterised by the presence of anti-nuclear antibodies (ANA) directed against entire nucleosomes and even naked DNA [32,33]. It is thought that the resulting immune complexes accumulate in joints, glomeruli and blood vessels manifesting as combined arthritis, glomerulonephritis (kidney disease) and vasculitis respectively. The precise cause of SLE is unknown although it has been suggested that the failure to adequately clear protein-DNA complexes liberated particularly during cell death may underlie the initiation and progression of this disease [33]. One of the many genes implicated in SLE is the DNA endonuclease DNase I (reviewed in [34]). DNase I represents the major nuclease activity present in serum, urine and other secreta. DNase I/- mice recapitulate the classical clinical spectrum of SLE including ANA production, deposition of immune complexes in glomeruli and overt glomerulonephritis [35]. Mutations in DNase I that significantly disrupt its enzymatic activity have also been found in other SLE patients [36]. Therefore, precedent for a functional association between compromised nuclease function, in this case a DNA endonuclease, and a systemic autoimmune disorder exists in the form of SLE. Interestingly, several monoallelic mutations in *Trex1* have recently been described in individuals with SLE and also in a variant of this condition, Familial Chilblain Lupus, which presents with ulcerating lesions of the skin in acral locations (fingers, ears and toes) [37,38].
9. Some outstanding questions.

The study by Yang, Lindahl and Barnes concerning the cellular consequences of TREX1-deficiency raises many questions that are likely to form the basis for further investigation [5]. For example, is the intracellular accumulation of ssDNA a cellular feature of non-\textit{Trex1}-mutated AGS? Moreover, if the ER-associated ssDNA that accumulates in the absence of TREX1 activity is derived from S-phase, how does it end up at the ER? Are replication forks particularly susceptible to collapse, even in the absence of exogenous DNA damage, in TREX1-deficient cells? If so, is it the collapse of these replication forks that ultimately triggers the DNA damage response? After all, TREX1-/- MEFs and TREX1-deficient AGS fibroblasts have a very low S-phase population [5]. DNA fibre technology will no doubt aid in addressing these issues.

Furthermore, p21 and p53 expression are clearly induced in the absence of TREX1 and consequently cells accumulate in G1. But, these cells can in fact cycle, albeit at a reduced rate [5]. This seems somewhat counter-intuitive if the chronically activated DNA damage-signalling pathway is fully functional. Yang \textit{et al} show that CHK2 protein levels are reduced, in an ATM-dependent manner, in the absence of TREX1 [5]. This suggests that the TREX1-deficient cells actively target a key effector of the DNA damage-signalling network to possibly de-sensitise the cell cycle checkpoint(s) facilitating some degree of growth. Interestingly, ATM levels in the AGS primary fibroblasts also appear to be reduced (Figure 5B [5]). Whilst chronic activation of the DNA damage response has been observed in other human disorders, for example Hutchinson-Gilford Progeroid syndrome, cells from these patients grow
extremely badly in culture [39,40]. The mechanism of the cell cycle progression in the absence of TREX1 activity awaits further characterisation at the molecular level.

One other aspect of the DNA damage response that is worthy of deeper investigation in the context of TREX1-deficiency (and RNase H deficiency) is the question of ATR’s contribution to the cellular phenotype observed by Yang et al [5]. After all, ssDNA is the lesion to which ATR responds [41,42]. Abundant evidence suggests that ATM is activated by dsDNA. It would be informative to examine ATM activation directly (eg. ATM-Ser-1981 autophosphorylation) in TREX1-deficient AGS cells to gain insight into the precise level of contribution of active ATM to the cellular phenotypes observed. It is becoming increasing evident that the interplay and functional overlap between ATM signalling and ATR signalling is greater than had previously been assumed. Complementary studies from the Jackson and Jeggo laboratories has provided evidence suggesting that ATM and ATR do not function as independent signalling pathways to distinct forms of DNA damage in human cells, but actually operate in concert [43,44]. There is increasing evidence that both CHK1 and CHK2 activities are required for a functional cell cycle arrest at G2 [43,44]. It is very interesting to note that whilst Yang et al do observe CHK1 phosphorylation following HU treatment in TREX1-/− MEFs, this signal is much reduced compared to wild type MEFs (Figure 3B [5]). This could be indicative of compromised ATR-pathway function in the absence of TREX1 or may simply reflect the low S-phase content of these cells. But, this response is reminiscent of the ‘leaky’ signalling responses observed in ATR-mutated Seckel syndrome cells [42,46]. It would be interesting to examine this in more detail using other ATR-pathway specific assays [45-47]. There is strong evidence to suggest that a compromised ATR-pathway is
associated with microcephaly in many human disorders [45-51], which provocatively, is frequently observed in AGS [2]. Whilst it is entirely speculative to suggest that this may potentially reflect compromised ATR-pathway function in AGS, it would be relatively straightforward to investigate this.

It remains to be determined why TREX1-/- mice do not reiterate the clinical phenotype of AGS, or conversely, why myocarditis is not a frequent clinical feature of AGS patients? Much remains to be uncovered about the cellular and clinical consequences of Trex1-deficiency. For example, novel monoallelic mutations in Trex1 have also been described as the cause of autosomal dominant Retinal Vasculopathy with Cerebral Leukodystrophy syndrome (RVCL) [52]. RVCL does not resemble AGS or SLE. Individuals with RVCL suffer visual loss, stroke and dementia. Interestingly, RVCL causative Trex1 mutations do not affect Trex1 exonuclease activity but impact on Trex1 intra-cellular localisation [52]. Work on TREX1 is currently uncovering novel interacting proteins and should provide more information on the multifunctional nature of this enzyme [16]. Finally, it will be important to uncover the extent of functional overlap, if any, between TREX1 and Trex2. Since further AGS causative genes await identification, it is conceivable that mutations in Trex2 could result in AGS (Table 1). I’m sure this question is actively being pursued.

10. The Future.

The story of TREX1 and AGS illustrates the power of the combined efforts of clinical geneticists, immunologists, biochemists and cell biologists in elucidating the molecular aetiology of a complex human disorder. If we can understand the molecular
origin of a disorder we can improve patient management and ultimately develop a curative therapy. A pertinent example is the use of DNase I-therapy in SLE (reviewed in [53]). Although initial optimism for the utility of this approach in SLE has been tempered by recent studies, nuclease treatment has been employed in the management of Cystic Fibrosis, a condition where the airways become blocked with viscous mucus containing very high concentrations of bacterial DNA (reviewed in [54]). TREX1-deficiency and AGS offer a fascinating example of why the cell must take extreme care in cleaning up its nucleic acid debris/metabolic intermediates, as it can lead to unexpected and unwelcome consequences. As is true for all ‘hot topics’ in research, much more information remains to be uncovered regarding Trex1, its physiological role and its contribution to the pathology of AGS.

Acknowledgements.

M O’D is a Cancer Research UK Senior Cancer Research Fellow. The M O’D laboratory is funded by CRUK and the UK Medical Research Council. Special thanks to P.A. Jeggo for critical reading of the manuscript.
Figure Legends.

Figure 1.

The structural domains of TREX1. Trex1 functions as a homodimer and its crystal structure has recently been solved [16,17]. For a detailed discussion of Trex1 structure-function analysis readers are referred to the studies by Brucet and colleagues and de Silva and colleagues [16,17]. Exo 1-3 represent the exonuclease domains which co-ordinate Mg$^{2+}$ which is essential for catalysis. PPII refers to the polyproline II motif implicated in mediating protein-protein interactions with proline-binding domain containing proteins. In fact, PPII has been shown to be required for the interaction of Trex1 with the transcription factor CA150 [16]. The functional significance of this interaction remains to be clarified. TMH is the transmembrane helix that traverses the ER membrane proposed to function in anchoring TREX1 to the ER [37]. Removal of the TMH is thought to be required for translocation of TREX1 to the nucleus. Trex1 mutations specifically affecting TREX1 intra-cellular localisation cause autosomal dominant Retinal Vasculopathy with Cerebral Leukodystrophy syndrome [52]. Aicardi-Goutieres syndrome, Systemic Lupus Erythematosus and Familial Chilblain Lupus causative mutations are distributed throughout the Trex1 gene. Work by de Silva and colleagues using TREX1 crystal structures illustrate how certain AGS causative Trex1 mutations affect TREX1 protein function at the molecular level [17].
Table 1.

A summary of the genes known to be involved in Aicardi-Goutieres syndrome.

<table>
<thead>
<tr>
<th>AGS locus mutation</th>
<th>Gene</th>
<th>Chromosome</th>
<th>Frequency of mutation in AGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGS1</td>
<td>Trex1</td>
<td>3p21.31</td>
<td>25%</td>
</tr>
<tr>
<td>AGS2</td>
<td>RNASEH2B</td>
<td>13q14.3</td>
<td>40%</td>
</tr>
<tr>
<td>AGS3</td>
<td>RNASEH2C</td>
<td>11q13.2</td>
<td>14%</td>
</tr>
<tr>
<td>AGS4</td>
<td>RNASEH2A</td>
<td>19p13.13</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No mutation identified</td>
</tr>
</tbody>
</table>

Table
References.


Figure 1