Non-sense mutations located in the 5' end of the coding sequence of a gene are commonly considered to be null alleles. Not only do such mutations result in the production of a truncated and usually inactive protein product, but premature stop codon mutations that occur upstream of the last exon–exon junction are also known to activate nonsense mediated decay (NMD) which results in the specific degradation of the affected mRNA. However, several mechanisms may allow either fully or partially functional protein to be produced from alleles containing a premature stop codon mutation and this phenomenon may lead to considerable amelioration of the resulting phenotype.

First, translational decoding (miscoding) of the stop codon as a sense codon, which can occur naturally due to the context of the stop codon, may result in low level "leaky" expression of full length protein. Although no examples of stop codon readthrough leading to disease amelioration have been clearly documented in humans, this mechanism has been implied for several premature stop codons, most notably for a premature stop codon mutation in the cystic fibrosis transmembrane conductance regulator gene (MIM 602421) leading to mild pulmonary presentation. In this case, the equivalent mutation in a yeast gene, Ste6, has been shown to be suppressed at levels as high as 10%. Second, an internally truncated form of the protein may be expressed due to altered processing of the mRNA in which the stop codon containing exon (and sometimes adjacent exons) are removed from the mature mRNA. A correlation between exon skipping and suppression of disease symptoms has been documented in several cases of Becker muscular dystrophy (MIM 300376), a milder form of Duchenne muscular dystrophy. Third, protein expression in some genes can be rescued by translational initiation at internal start codons downstream from a premature stop codon or frameshift mutation. This might occur by leaky scanning, internal ribosome entry, or through reinitiation following termination of translation. This kind of mechanism has been implicated for mild disease presentation in several diseases, including adrenal hypoplasia congenita (AHC; MIM 300200), Nijmegen breakage syndrome (NBS; MIM 251260), attenuated adenomatous polyposis coli (APC; MIM 175100), and one form of peroxisome-biogenesis disorder (PBD; MIM 601758).

We examined the possibility that one of these mechanisms might be responsible for the unusually mild phenotype of patients carrying a premature stop mutation in the ATRX gene. ATRX is an X encoded member of the SNF2 protein family which exhibits ATPase and chromatin remodelling activities. Constitutional mutations give rise to a form of syndromal mental retardation (ATR-X syndrome; MIM 300032) associated with severe to profound learning difficulties, facial dysmorphism, genital abnormalities, and alpha thalassaemia. Somatic mutations have recently been identified in the clonal haematological disorder alpha thalassaemia myelodysplasia syndrome (ATMDS; MIM 300448). In ATR-X syndrome, the majority of mutations are missense and lie within two highly conserved regions of the gene, a PHD-like zinc finger domain and a helicase domain. True null mutations appear to have severe consequences: a mouse ATRX knockout is embryonic lethal (unpublished data). It was intriguing, therefore, to find that a milder variant of ATR-X syndrome was associated with a 109C→T mutation which gives rise to a premature stop codon, R37X. In the family reported by Guerrini et al, of the four affected male cousins, one had profound mental retardation which is typical of the condition, but a second had moderate mental retardation and two had mild mental retardation and were able to live partially independent lives. Subsequently, another unrelated male with moderate mental retardation has been identified with the same mutation.

In this report we demonstrate, through analysis of the endogenous protein, that cell lines from individuals with the R37X mutation express approximately 20% of the wildtype levels of ATRX protein. Alternate splicing is excluded as the mechanism for this level of protein expression. We show that it is not due to stop codon readthrough, but instead it is

**Key points**

- Premature stop codon mutations that occur near the 5' end of the coding sequence of a gene are commonly considered to be null alleles. However, further characterisation of patients in whom the observed phenotype is much milder than expected for a null mutation has revealed several mechanisms by which functional protein can be produced and phenotypic rescue achieved.
- We demonstrate in vivo, a mechanism invoking initiation of translation at a site downstream of a potentially lethal premature termination mutation, R37X, in the X linked ATRX gene (MIM 300032). The resulting phenotype can be very mild and is associated with levels of ATRX protein that are up to 20% of those seen in the wildtype.
- This phenomenon may be an overlooked mechanism to explain other situations in which unexpectedly mild phenotypes arise from nonsense mutations that occur in the 5' coding region of genes.

**Abbreviations:** AHC, adrenal hypoplasia congenita; APC, adenomatous polyposis coli; ATMDS, alpha thalassaemia myelodysplasia syndrome; NBS, Nijmegen breakage syndrome; NMD, nonsense mediated decay; PBD, peroxisome-biogenesis disorder
associated with translation initiation at an in-frame internal AUG located three codons downstream of the R37X stop codon mutation. Phenotypic rescue due to the persistent expression of amino truncated ATRX protein from the downstream initiation site may provide an explanation for the mild presentation of R37X patients. This unusual mechanism, by which a null mutation may be “rescued” by translational initiation at a downstream start site, may underlie other inconsistencies in relating genotype to phenotype in human genetic disease.

METHODS

Patients
The study was approved by the Multi-centre Research Ethics Committee (ref MREC 02/01/02). The family in which the R37X mutation was originally identified has been previously described (cases 1-4 in Guerrini et al16).

RNA and protein analysis
Epstein-Barr virus transformed lymphoblastoid cell lines were used as a source of RNA and nuclear protein extract.

A 2 µg sample of total RNA from each cell line was reverse transcribed using the ProSTAR first strand RT-PCR kit (Stratagene, La Jolla, CA, USA). The generated first strand cDNA was diluted and used as template for real time quantitative PCR analysis (TaqMan, Applied Biosystems, Foster City, CA, USA). Assays specific for ATRX (Hs00230877_m1) and GAPD (Hs99999905_m1) mRNAs were obtained from Assays-on-demand (Applied Biosystems). Each sample was analysed in triplicate and the increase of the fluorescence at each PCR cycle was then measured following injection of 50 µl of either Renilla or firefly luciferase substrate. Luciferase activity was determined using the Dual Luciferase Reporter Assay System (Promega). Relative light units were measured on an MLX microplate luminometer (Dynex Technologies Chantilly, VA, USA). Transfected cells were lysed in 12.5 µl passive lysis buffer and light emission was measured following injection of 50 µl of either Renilla or firefly luciferase substrate.

Luciferase reporter constructs
Oligonucleotides were synthesised at the University of Utah Core Facility. Oligonucleotides were designed such that by a two step PCR amplification, the 5’ end of the ATRX mRNA with the appropriate restriction sites was produced for cloning into phRL-CMV (Promega, Madison, WI, USA). The region of amplification begins at the 5’ end of the mDNA and extends 134 nucleotides downstream of the first AUG of the coding region. The wildtype ATRX region was PCR amplified using Taq DNA Polymerase (Sigma-Aidrich, St Laws, MO, USA) in a Perkin Elmer 9600 thermocycler using human genomic DNA (Promega) as a template. R37X and M40L changes were made by designing single nucleotide changes into the 3’ primer to generate PCR products containing the respective single nucleotide changes. The PCR products were digested with the appropriate restriction enzymes for ligation into phRLuc. Each construct was sequence verified at the University of Utah Core Facility.

Cell culture and transfections
The human embryonic kidney cell line, HEK293, was obtained from ATCC and maintained as previously described13 in the absence of antibiotics. Cells used in these studies were subcultured at 70% confluence and used between passages 7 and 15. Cells were transfected using Lipofectamine 2000 reagent (Invitrogen, Paisley, UK), using the one-day protocol in which suspension cells are added directly to the DNA complexes in 96 well plates. A 25 ng aliquot of DNA and 0.2 µl Lipofectamine 2000/well in 25 µl Opti-Mem (Gibco, Invitrogen) were incubated and plated in opaque 96 well half area plates (Costar, Corning, NY, USA). Cells were trypsinised, washed, and added at a concentration of 4×10^4 cells/well in 50 µl DMEM, 10% FBS. Transfections were incubated overnight at 37°C in 5% CO2, then 75 µl DMEM, 10% FBS was added and incubation continued for 48 h.

Dual luciferase assay of stop codon readthrough
Luciferase activity was determined using the Dual Luciferase Reporter Assay System (Promega). Relative light units were measured on an MLX microplate luminometer (Dynex Technologies Chantilly, VA, USA). Transfected cells were lysed in 12.5 µl passive lysis buffer and light emission was measured following injection of 50 µl of either Renilla or firefly luciferase substrate.

![Western blot analysis of lymphoblastoid cell line from an individual with the R37X mutation. Nuclear extracts were probed with the anti-ATRX monoclonal 39f. Lanes: N, normal individual; R37X, an affected individual. Sample loading was adjusted to give equivalent signals in normal and R37X lanes. In the normal lane, three ATRX specific bands are seen (of ~300 kDa and a doublet of ~200 kDa); in R37X each of these bands is slightly reduced in size.](http://www.jmg.bmj.com/)

Figure 1 Western blot analysis of lymphoblastoid cell line from an individual with the R37X mutation. Nuclear extracts were probed with the anti-ATRX monoclonal 39f. Lanes: N, normal individual; R37X, an affected individual. Sample loading was adjusted to give equivalent signals in normal and R37X lanes. In the normal lane, three ATRX specific bands are seen (of ~300 kDa and a doublet of ~200 kDa); in R37X each of these bands is slightly reduced in size.

www.jmedgenet.com
RESULTS

Translation of ATRX carrying the R37X mutation in vivo

Western blot analysis of ATRX protein was carried out using a monoclonal antibody whose epitope is downstream of the predicted premature stop codon. Analysis of nuclear extracts derived from patient cells carrying the R37X mutation revealed the presence of ATRX protein which is fractionally smaller than that seen in normal controls (fig 1). At least two ATRX protein isoforms are normally present (300 and 200 kDa) and it is evident that all of these are affected in patients with the R37X mutation. Levels were reduced as compared with a normal control but comparison of signal strengths indicated that cells carrying the R37X mutation expressed approximately 20% of wildtype levels of ATRX (fig 2A and B). In the lymphoblastoid cells examined, there was a degree of variability in the ATRX protein levels between affected individuals, but there was no obvious correlation between this and their intelligence quotient (data not shown). There were comparable levels of ATRX mRNA in wildtype and mutant cells as measured by quantitative PCR showing that R37X ATRX mRNA is resistant to NMD in these cells (fig 3). Amplification of cDNA from normal and affected individuals using primers in exons 1 and 4 showed no detectable alternate splicing which might lead to skipping of the mutation in exon 2 (fig 4).

Translational initiation in the ATRX gene in vitro

Examination of the 5' end of the ATRX gene revealed the presence of three AUG codons downstream from the previously described initiation codon. Comparison of the potential ATRX initiation codon sequence contexts to the most highly conserved nucleotide positions known to facilitate translation initiation (−3 (A or G) and +4 (G)) revealed that while all AUG codons deviated at the +4 position, the AUGs at codon position 1 and 40 contained a purine match at the important −3 position (table 1). The presence of a potential start codon downstream of the R37X mutation suggested that expression of ATRX protein via an alternate initiation event might be responsible for the observed expression of a slightly shortened form of the protein in samples derived from the ATR-X patients.

In order to determine if alternative translation initiation at the AUG codon downstream of the premature stop codon mutation could be utilised for initiation, the 5' end of the ATRX gene, the R37X allele, and each allele with the downstream AUG codon changed to a CUG codon were cloned upstream of a Renilla luciferase reporter gene in phRL-CMV resulting in plasmid ATRX-Rluc, ATRX M40L-Rluc, R37X-Rluc, and R37X M40L-Rluc, respectively (fig 5). In vitro transcription and translation of ATRX-Rluc using T7 RNA polymerase and rabbit reticulocyte lysates revealed two protein products differing by approximately 4 kDa (fig 6A, lane 1). Changing the AUG 40 methionine codon to a CUG leucine codon, ATRXM40L, eliminated the smaller band demonstrating that this product arises from initiation at codon 40. In R37X-Rluc the product of in vitro transcription and translation was exclusively the smaller product, and likewise, changing the AUG at codon 40 to a CUG, R37XM40L, eliminated this protein band (fig 6A, lanes 3 and 4). Finally, the lack of full length ATRX protein in the
context of the R37X mutation, in vitro and in vivo, indicates that stop codon readthrough occurs at a low level and does not contribute significantly to ATRX expression. It appears that in vitro, translation can initiate at two AUG codons, one located upstream of the R37X mutation at the previously described initiation codon and one downstream at codon 40.

The efficiency of Renilla luciferase expression was also examined in tissue culture cells transfected with ATRX-Rluc, ATRXM40L-Rluc, R37X-Rluc, or R37XM40L-Rluc. Consistent with utilisation of the downstream methionine at codon 40, R37X-Rluc maintained approximately 15% Renilla luciferase activity relative to the ATRX construct and changing the AUG at position 40, R37XM40L-Rluc, resulted in background levels of luciferase expression (fig 6B).

In vivo, in the wildtype, a single product is observed consistent with translation initiation at codon 1. In the presence of the R37X mutation, a single shorter protein product is observed. The reduced size of this product is consistent with data from both in vitro translations and transfection of cultured cells, which indicates that translation can initiate downstream of the premature stop codon mutation at the start codon in position 40 resulting in an amino truncated ATRX protein.

**DISCUSSION**

ATRX appears to play a vital role in regulating gene expression. In its absence, mouse development arrests and the embryo fails to develop beyond 9.5 days (unpublished data). The mild phenotype associated with the R37X mutation was clearly inconsistent with this observation and here we have shown that affected individuals are able to make substantial amounts of ATRX protein through the use of a second, efficient translation initiation site located at codon 40 in exon 2. In the absence of a functional assay for ATRX it is not possible to study the consequence of losing the N terminal 39 amino acids although we have shown that the pattern of nuclear localisation is not perturbed (data not shown). Since the degree of intellectual handicap in some affected individuals is relatively mild, it is possible that the function of the shortened protein is preserved and that the variation in severity between individuals reflects the efficiency of initiation at M40 in brain. In lymphoblastoid cells, no significant differences in ATRX protein levels could be detected between individuals of different intelligent quotients. One possibility is that, functionally important differences in protein expression are too subtle to be distinguished by western blot. An alternative explanation is that the levels of mRNA or efficiency of downstream initiation in different individuals is not consistent in all tissues. A recent paper has demonstrated that collagen X transcripts containing a premature stop codon are subject to NMD in a tissue specific manner. And, certainly in ATR-X syndrome, there is no correlation in the severity of the phenotype in one organ system and the severity in another—an individual with profound mental retardation will not necessarily have a severe form of alpha thalassaemia.

R37X message levels are only slightly reduced in lymphocytes carrying the mutation indicating a resistance to NMD in this cell line. One possible explanation for this observation is that a low level of translation of the message downstream from the stop codon is sufficient to protect the message from NMD. A previous report of premature stop codons near the 5’ end of the triosphosphate isomerase gene (MIM 190450) indicates that reinitiation of translation at a downstream start codon can abrogate nonsense mediated mRNA decay of this mutant message. The abundance of R37X containing message and the lack of detectable altered splicing products is strong evidence that the foreshortened protein product is derived from normally processed mRNA containing the premature stop codon.
Values were normalised to a co-transfected firefly reporter plasmid. Renilla luciferase activity produced relative to ATRX-Rluc is shown. All HEK293 cells. For ATRXM40L-Rluc, R37X-Rluc, and R37XM40L-Rluc, the R37X-Rluc, and R37XM40L-Rluc were transiently transfected into indicated by arrows: M1 and M40. (B) ATRX-Rluc, ATRXM40L-Rluc, bands from initiation at the first and 40th ATRX methionine codon are translation downstream of a disease causing mutation. While both bypass upstream initiation codon(s) may also result in initiation near the 5'9 initiation at the downstream start codon occurs in the context of the premature stop mutation and allows remarkably high levels of ATRX protein to be expressed. Bypassing of the premature stop codon in this X linked condition was recognised because just a single allele was present. It is possible that in the context of deleterious mutations, initiation at downstream codons is an important but under-recognised phenomenon which gives rise to unexpectedly mild phenotypes. The amelioration of disease may lead to diagnostic confusion. Furthermore, the aberrant genotype-phenotype correlation that results from this phenotypic rescue may lead to a misunderstanding of protein function and disease pathophysiology.

In wildtype cells, only protein initiated at M1 is detectable by western blot whereas in cells carrying the R37X mutation initiation is from M40, although the method of detection cannot exclude the presence of a low level of the foremost protein from the wildtype. In vitro transcription translation reactions demonstrate that both start codons may be used for initiation in the wildtype. While it is clear that initiation at the downstream start codon occurs in the context of the R37X mutation with high efficiency both in vitro and in vivo, additional experiments are required to fully elucidate the mechanism of initiation in this context. Initiation at an internal start codon can occur by several mechanisms, such as termination dependent reinitiation, leaky scanning, or internal ribosome entry. For termination dependent reinitiation, ribosomes encountering a stop codon mutation near the 5’ end of a gene may remain associated with the message following termination and reinitiate at a downstream start codon. Likewise, leaky scanning in which some ribosomes bypass upstream initiation codon(s) may also result in initiation downstream of a disease causing mutation. While both termination dependent reinitiation and leaky scanning depend upon ribosomes that enter the message through the 5’ cap and occur with greatest efficiency near the 5’ end of the message, internal ribosome entry at an internal ribosome entry site may occur anywhere along the message. In each case, if the downstream start codon is in frame and utilised efficiently, as observed with the R37X mutation, partial phenotypic suppression of the mutation may occur.

It is not known how frequently phenotypic rescue of deleterious mutations occurs through the use of a second internal translation initiation site but evidence is accumulating to suggest that this is an important phenomenon. Translation initiation at an internal start codon has been proposed as an explanation for a mild form of AHC and PBD arising from a premature stop codon in the Dax1 gene and a truncating frameshift mutation in the PEX12 gene, respectively. Also, mutations in the 5’ region of the adenomatous polyposis coli tumour suppressor gene are responsible for a milder form of adenomatous polyposis coli, attenuated APC, which has been explained by the presence of an internal ribosome entry site that directs initiation to a downstream start site. In the Nijmegen breakage syndrome, internal translational initiation appears to give rise to a truncated and partially functional protein associated with the 657del5 frameshifting mutation.

For the ATRX mutation R37X, we have examined phenotypic rescue through both in vitro and in vivo studies. Both results show that initiation at the downstream start codon occurs in the context of the premature stop mutation and allows remarkably high levels of ATRX protein to be expressed. Bypassing of the premature stop codon in this X linked condition was recognised because just a single allele was present. It is possible that in the context of deleterious mutations, initiation at downstream codons is an important but under-recognised phenomenon which gives rise to unexpectedly mild phenotypes. The amelioration of disease may lead to diagnostic confusion. Furthermore, the aberrant genotype-phenotype correlation that results from this phenotypic rescue may lead to a misunderstanding of protein function and disease pathophysiology.

Authors’ affiliations
M T Howard, C B Anderson, J F Atkins, Department of Human Genetics, University of Utah, Salt Lake City, UT 84112-5330, USA
N Malik, J L A Voskuil, R J Gibbons, MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, OX3 9DS, UK
This work was supported by MRC (RG, NM, JV) and NIH GM48152 (JFA) and the Muscular Dystrophy Association (JFA and MTH).
Conflict of interest: none declared.
Correspondence to: Dr Richard Gibbons, MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, OX3 9DS, UK; richard.gibbons@imm.ox.ac.uk
Revised version received 13 May 2004
Accepted for publication 20 May 2004

ELECTRONIC-DATABASE INFORMATION

REFERENCES


