A complex deletion-inversion-deletion event results in a chimeric IL1RAPL1-dystrophin transcript and a contiguous gene deletion syndrome

J M Wheway, S C Yau, V Nihalani, D Ellis, M Irving, M Splitt, R G Roberts

We originally identified the IL1RAPL1 gene through its partial deletion in a patient with Becker muscular dystrophy (BMD), glycerol kinase deficiency (GKD), adrenal hypoplasia congenita (AHC), and mild mental retardation, and suggested that its disruption might account for the patient’s cognitive problems. Other workers have shown that intragenic mutations of the IL1RAPL1 gene are associated with mild, non-specific X linked mental retardation. IL1RAPL1 encodes a putative transmembrane cytokine receptor related to receptors and receptor accessory proteins for interleukin-1 and -18, and is expressed in brain and muscle. The protein structure comprises three extracellular immunoglobulin domains, which presumably mediate binding of an as yet unidentified ligand, a transmembrane region, and an intracellular domain, which is likely to enable signalling via the NFκB pathway. Its role in brain development or function is not understood.

We describe here the precise structure of a complex deletion-inversion-deletion mutation involving more than 2 Mb of Xp21 in a patient with a contiguous gene deletion syndrome. This results in the loss of several genes and the creation of an unusual fusion of the IL1RAPL1 and dystrophin genes with the potential to generate a chimeric protein product. We discuss the possible mutational mechanisms and biological consequences of this rearrangement.

MATERIAL AND METHODS

Clinical description
The patient was the second child of healthy, non-consanguineous parents (father 39 years old, mother 17). He was well in the immediate neonatal period, but collapsed at 17 hours with profound hypoglycaemia. In spite of intravenous dextrose, he had a cardiopulmonary arrest. Resuscitation required several doses of intravenous epinephrine and he eventually responded to intravenous hydrocortisone. Subsequently, he required artificial ventilation for four days. Following resuscitation, he was found to be hyponatraemic and ultrasound scanning failed to identify the adrenal glands. Further investigation showed glycerol kinase deficiency (0.3 mmol/h/mg protein in cultured skin fibroblasts), an undetectable random cortisol level, occipital changes on MRI scanning consistent with hypoglycaemic and hypoxic-ischaemic insults, absent visual evoked responses, and raised serum creatinine kinase (peak value 8703 U/l). Routine chromosome analysis showed a 46,XY normal male karyotype. At 10 months, he had not had any metabolic complications, but was blind with microcephaly (<0.4th centile) and had significant motor delay secondary to hypoxic-ischaemic/hypoglycaemic cerebral damage. He was feeding well and thriving on a low fat diet with high calorie supplements, and was receiving daily fludrocortisone replacement therapy.

Key points
- We describe here an unusual case of complex GKD in which a deletion of ∼1.6 Mb has removed the GKD and AHC genes, together with part of the mental retardation gene IL1RAPL1, the MAGEB gene cluster, and the tests specific ferritin heavy chain gene FTHL17. A deletion of 35 kb has removed exon 52 of the dystrophin gene. The intervening 600 kb region, containing exons 53–59 of the dystrophin gene, is inverted.
- This deletion-inversion-deletion results in a chimeric IL1RAPL1-dystrophin transcript capable of encoding a hybrid protein.
- Two of the four breakpoints were found to lie in Alu repeats, and the rearrangement seems to have occurred during grandpaternal meiosis, implicating inappropriate recombination within the unpaired X chromosome.

Reverse transcript PCR
RT-PCR was performed as described previously; sequences of supplementary primers are available in Supplementary information (www.jmedgenet.com).

Genomic PCR
For genomic PCR reactions, the following conditions were used. Where the special conditions used for long PCRs across breakpoints differ, these are given in parentheses. Twenty-five ng of genomic DNA were amplified in 25 µl using 2.5 units of Promega Taq polymerase (1.75 units of Roche High Fidelity Taq polymerase) in 1 x manufacturer’s buffer, 2.5 mmol/l MgCl₂, 0.5 mmol/l dNTPs, and 10 ng of each primer. PCR was performed in a Progene Techne thermocycler (MJ Research Dyad thermocycler) using the following parameters: 94°C for five minutes followed by 30 cycles of 94°C for one minute, 58°C for one minute, 72°C for two minutes, followed by 72°C for five minutes (94°C for two minutes followed by 30 cycles of 94°C for 15 seconds, 60°C for 30 seconds, 68°C for three minutes extended by five seconds per cycle, followed by 68°C for five minutes). Primer sequences are available in Supplementary information.

3’-RACE with vectorette
3’ rapid amplification of cDNA ends (3’-RACE) with vectorette was performed as described previously, using 500 ng of patient total RNA and gene specific primers lying in exons 46 and 49 of the dystrophin gene (see Supplementary information). The product was cloned into pCRII-TOPO (Invitrogen) and sequenced.

Sequence analysis
Genomic PCR products spanning the breakpoints and the IL1RAPL1-dystrophin RT-PCR product were analysed using an
ABI 377 automated fluorescent sequencer according to the manufacturer's instructions.

RESULTS

Delineation of deletions

The syndrome of DMD (or BMD), GKD, and AHC is invariably associated with a substantial deletion in Xp21, removing varying amounts of the 3′ end of the dystrophin gene together with the DAX1 and GK genes (fig 1A). However, routine deletion analysis of the proband's dystrophin gene by multiplexed fluorescent PCR unexpectedly showed the isolated loss of exon 52, with preservation of subsequent exons. Although the deletion of the frameshifting exon 52 is consistent with a DMD phenotype, the retention of the 3′ end of the gene was surprising given the apparent contiguous gene deletion syndrome. Moreover, amplification using primer pairs at each end of the GK gene confirmed that this gene was deleted in its entirety.

The existence of two non-contiguous deletions in this patient initially seemed to suggest that the DMD and GKD/AHC might have arisen through two separate mutational events, but clearly there remained the possibility that the two deletions were part of a single, more complex rearrangement. Indeed, attempted RT-PCR of the dystrophin transcript in overlapping sections yielded products from reactions both 5′ and 3′ of exon 52, but no reaction which traversed exons 51-53 amplified successfully. This situation was reminiscent of that observed during attempted amplification of factor VIII transcripts from patients bearing large inversions.

In order to clarify the situation, we first set out to determine the limits of both deletions by a process of reiteratively dividing the regions known to contain the breakpoints. Introns 51 and 52 of the dystrophin gene (45 and 50 kb, respectively) were each divided into four sections, bounded by newly designed sequence tagged sites (STSs). After testing the STSs on the patient, the section between the first missing STS and last retained STS was in turn divided into four and the process repeated until breakpoints 3 and 4 were pinpointed to within 1 or 2 kb (figs 1B and Supplementary information).

A large scale STS based deletion screen of the ∼2 Mb telomeric to the dystrophin gene showed that while the last exon (exon 79) of the dystrophin was retained, the testis specific ferritin heavy chain gene (FTHL17), a mere 47 kb away, was deleted. At the telomeric end of the deletion, although exons 1-5 of the IL1RAPL1 gene were present, exons 6-11 were missing. These two breakpoints (designated 1 and 2 respectively, fig 1B) were sublocalised within the 264 kb IL1RAPL1 intron 5

Figure 1 Structure of the rearrangement. (A) Diagram of the region of Xp21 between the IL1RAPL1 and dystrophin genes. The better characterised genes in the intervening region are shown, together with the extents of known mutations which result in fusions of the two large flanking genes. (B) The structure of the rearranged genomic region in the patient. 1, 2, 3, and 4 indicate the rearrangement breakpoints. Shaded boxes indicate genes and vertical lines are their exons. Both [A] and [B] are drawn to scale from the current draft human genome sequence. (C, D) Wild type (WT) sequences of the breakpoints are aligned with chimeric sequences across the patient's rearrangement breakpoints. (C) Junction 1:3 between IL1RAPL1 intron 5 and dystrophin intron 52 (accession number AF533680). Light box indicates Alu repeats. Bold box indicates the site of the junction. (D) Junction 2:4 between the FTHL17-dystrophin intergenic region and dystrophin intron 51 (accession number AF533681). Bold box indicates the novel inserted sequence at the junction.
and the 47 kb FTBL17-DMD interval by the process of reiterative division described above.

The above experiments showed that a 35 kb region of the dystrophin gene was missing from the patient, comprising the 20 kb before exon 52 and the 15 kb after. In addition, a large region (∼1.6 Mb) from 60 kb into intron 5 of the IL1RAPL1 gene to 9 kb outside the dystrophin gene has been lost (fig 1B).

Molecular characterisation of rearrangement

As we had defined each deletion breakpoint to within a kilobase or two, it should now have been possible to amplify across the breakpoints in order to characterise the rearrangement(s) at the nucleotide level. Because the gross nature of the rearrangement was unknown, we attempted amplification using every pairwise combination of primers from the STSs closest to each of the four breakpoints. Only two of the six possible combinations yielded PCR products from the patient sample, namely that between breakpoints 1 and 3, and that between breakpoints 2 and 4.

Sequence analysis of these products (fig 1C, D) showed them to be bona fide chimeras of the expected genomic sequences. The junction between breakpoints 1 and 3 (accession number AF533680) is clean and occurs just within the end of Alu repeats in both parent sequences (a left hand half Alu at breakpoint 1, and an intact Alu at breakpoint 3). The Alu repeats confer ∼80% identity on the regions containing the two breakpoints and are almost certainly responsible for orchestrating the rearrangement. The junction between breakpoints 2 and 4 (accession number AF533681) shows no appreciable similarity between the two wild type sequences, and bears 15 novel base pairs which are not present in either parent sequence.

The most parsimonious interpretation of these results is that the 600 kb region between the two deletions, which contains exons 53-79 of the dystrophin gene, is inverted in our patient. The Alu repeats confer ∼80% identity on the regions containing the two breakpoints and are almost certainly responsible for orchestrating the rearrangement. The junction between breakpoints 2 and 4 (accession number AF533681) shows no appreciable similarity between the two wild type sequences, and bears 15 novel base pairs which are not present in either parent sequence.

The most parsimonious interpretation of these results is that the 600 kb region between the two deletions, which contains exons 53-79 of the dystrophin gene, is inverted in our patient. This finding enabled two predictions to be made regarding the transcriptional consequences of the deletion-inversion-deletion mutation. First, transcription of the IL1RAPL1 gene should pass from intron 5 into the sense strand of intron 52 of the dystrophin gene. Exon 5 of IL1RAPL1 is therefore expected to splice onto exon 53 of the inverted 3′ 600 kb of the dystrophin gene, yielding a chimeric transcript. Second, transcription of the dystrophin gene should traverse intron 51 and enter the antisense strand of the 3′ end of the same gene. It is then likely to find a cryptic 3′ exon and terminate.

Figure 2  Consequences, origin, and postulated mechanism of the mutation. (A) Postulated protein product of the IL1RAPL1-dystrophin chimeric transcript, showing fusion of the first two immunoglobulin domains of IL1RAPL1 to the C-terminal portion of dystrophin. (B) Sequence across the junction of the RT-PCR product obtained from the patient, together with its notional translation. (C) Results of genotype analysis at multiplexed microsatellite markers across Xp21-q22. The arrow indicates the dystrophin gene. The markers are (from bottom to top): DYSI, DYSII (just 5′ of the dystrophin gene), S5N3, STR44, STR45, STR49, DXS1214 (introns 1, 44, 45, 49, and 63 of the dystrophin gene, respectively), DXS989, DXS1202 (4 and 6 Mb 3′ of the dystrophin gene). (D-G) Postulated mechanism for the deletion-inversion-deletion mutational event. (D) Wild type X chromosome showing breakpoints 1, 2, 3, and 4. (E) Proposed aberrant recombination events during grandpaternal meiosis. (F) Resolution of recombinations to give the deletion-inversion-deletion product and a circular molecule containing both deleted regions. (G) The circular product is lost and the rearranged chromosome is inherited by the patient’s mother.
Charaterisation of the chimeric transcripts and proteins

In order to test the first prediction that a chimeric mRNA is formed from the fused IL1RAPL1 and dystrophin genes in our patient, we performed nested RT-PCR on RNA from muscle biopsy material using forward primers in exon 1 of the IL1RAPL1 gene and reverse primers in exon 58 of the dystrophin gene. This yielded large amounts of a single product of 1806 bp, the size expected for E5-E53 splicing. Sequencing of this product showed the predicted in-frame splicing (accession number AF533682), which would be expected to direct translation of an IL1RAPL1-dystrophin fusion protein (fig 2A, B).

With a view to testing the second prediction, we also attempted 3′-RACE with vector ette, using nested primers in exons 46 and 49 of the dystrophin gene. This yielded a single product of 658 bp, comprising the 3′ end of exon 49 spliced to exons 50 and 51, followed by the 215 bp of intron 51 which lie immediately adjacent to exon 51 in the genome. At this point, a run of A residues follows a consensus AUUAAA polyadenylation signal.

Both transcripts are expected to be resistant to nonsense mediated decay, as they have open reading frames which proceed uninterrupted into the 3′ most exon. However, we had insufficient patient RNA to test this empirically by comparing dystrophin and IL1RAPL1 mRNA levels with those of normal muscle. Anecdotally, however, amounts of RT-PCR products obtained were of the same order as those obtained from normal controls.

Although no antibodies are available for IL1RAPL1, the existence of antibodies against different regions of the dystrophin protein enabled us to monitor both of the proposed protein products of the inversion. Antibodies DYS3 and DYS1 (against amino acids encoded by exons 10-12 and 26-30, respectively; Novocastra Laboratories Ltd, Newcastle upon Tyne, UK) should detect the truncated E1-E51 product, while DYS2 (exons 78-79) should detect the IL1RAPL1-dystrophin E53-E79 fusion protein. Negligible signal was seen when any of these three antibodies was used for immunofluorescent staining of slices of muscle biopsy material from our patient (data not shown; anti-spectrin antibodies were used as a positive control for quality of the preparation). We deduce from this that neither of the protein products is significantly stable.

Origin of mutation

In order to determine the origin of mutation in this family, we assessed genotypes at a range of polymorphic markers spanning a region from just 5′ of the dystrophin gene (DYS1) to approximately 6 Mb 3′ of the dystrophin gene (DXS989, beyond IL1RAPL1). This showed (fig 2C) that the region of the X chromosome rearranged in the patient is of grandpaternal origin. Quantitative fluorescent multiplex PCR analysis of the dystrophin gene showed that while the patient’s mother carries the mutation, his maternal aunt, who also shares this haplotype, does not.

DISCUSSION

Consequences of deletion

The two deletions described here are individually relatively unremarkable. Intragenic deletions of one or more exons of the dystrophin gene account for some two-thirds of cases of DMD, while many instances of contiguous gene deletion involving this region have been reported. The region from the 5′ end of the IL1RAPL1 gene to the 3′ end of the chronic granulomatous disease gene (including the dystrophin gene) represents a >9 Mb stretch (>5% of the X chromosome) whose complete loss is still compatible with life; this is a substantial mutational target. Mutations which individually affect the IL1RAPL1, DAX-1, GK, and dystrophin genes have been shown to result in mild learning difficulties, AHC, and BMD/DMD, respectively. The phenotypic consequences of deleting the MAGE family or the testis specific ferritin heavy chain gene FTTH17 (or the many other less well characterised coding regions) are as yet unknown, and likely to be masked in this severely disabled patient. Mutations in a chromosome 19 linked ferritin light chain gene are associated with a Huntington-like neurodegenerative disorder.

Consequences of inversion

The inversion does not result per se in any further disruption of the genes in the region. Instead it affords the opportunity for a unique chimeric IL1RAPL1-dystrophin transcript to arise. The principal of expression of IL1RAPL1 are skeletal muscle and brain; it is therefore expected that these should be main sites of expression of the chimeric transcript in our patient. We expected the transcript to be relatively stable, as it encodes a continuous open reading frame and possesses a full length dystrophin 3′UTR. Accordingly, we were able to amplify the chimeric transcript from the patient’s skeletal muscle RNA with ease. Its translation would give rise to a 1366 residue protein (fig 2A) comprising the first two of the three immunoglobulin domains of IL1RAPL1 (amino acids 1-234) followed by the last four spectrin repeats and the C-terminal domains of dystrophin (amino acids 2555-3686). The presence of a signal sequence at the N-terminus of IL1RAPL1 makes it likely that the chimeric protein is inserted into the endoplasmic reticulum (ER), where it would either be glycosylated and secreted or proteolysed. It seems probable that the oxidising milieu of the ER would cause gross misfolding of the normally cytoplasmic, cysteine rich dystrophin C-terminus, with consequent proteolytic degradation.

Consistent with this, no signal was detected in the patient’s skeletal muscle using the DYS2 antibody against the C-terminal 17 amino acids of dystrophin.

The inversion of sequences 3′ to dystrophin exon 51 is expected to result in a dystrophin protein lacking the C-terminal membrane binding domains. Consistent with this, we performed nested RT-PCR on RNA from muscle biopsy material with ease. Its translation would give rise to a 1366 residue protein (fig 2A) comprising the first two of the three immunoglobulin domains of IL1RAPL1 (amino acids 1-234) followed by the last four spectrin repeats and the C-terminal domains of dystrophin (amino acids 2555-3686). The presence of a signal sequence at the N-terminus of IL1RAPL1 makes it likely that the chimeric protein is inserted into the endoplasmic reticulum (ER), where it would either be glycosylated and secreted or proteolysed. It seems probable that the oxidising milieu of the ER would cause gross misfolding of the normally cytoplasmic, cysteine rich dystrophin C-terminus, with consequent proteolytic degradation.

Consistent with this, no signal was detected in the patient’s skeletal muscle using the DYS2 antibody against the C-terminal 17 amino acids of dystrophin.

The inversion of sequences 3′ to dystrophin exon 51 is expected to result in a dystrophin protein lacking the C-terminal membrane binding domains. Consistent with this, no signal was detected in the patient’s skeletal muscle using the DYS2 antibody against the C-terminal 17 amino acids of dystrophin.

Inversion mutations in human disease

Probably because of their difficulty of detection, there are not many published reports of subcytogenetic inversions. Despite their ability to bisect a gene, they do not result in quantitative changes in the inverted area, and are only detectable by their junction fragments on a Southern blot or (in the case of X linked diseases) by the failure of PCR reactions which traverse the breakpoints. The prohibitive expense of designing PCR reactions to traverse any possible inversion breakpoint means that most genomic mutation screening strategies are incapable of detecting pathogenic inversions. Because RF-PCR of transcripts is sensitive to long range integrity of the gene, however, transcript based screening strategies of X linked genes can readily detect the presence of an inversion. Hence, the common factor VIII inversions responsible for 50% of cases of severe haemophilia A were initially discovered through their effect on attempted RF-PCR traversing the boundary between exons 22 and 23 of that gene. The disruption of RF-PCR also contributed to our seeking an inversion in the patient described here. Clearly, the de novo detection of inversions in autosomal genes remains challenging.

Mutational mechanism

Examination of the relatives of the affected boy indicated that the deletions and the inversion probably occurred as a single concerted event during grandpaternal gametogenesis (fig 2C). Previous work on high frequency (4×107 per gene per gamete per generation) inversion events in Xq28 which result in
severe haemophilia A has shown that these too occur overwhelmingly (69/70 events) during male gametogenesis. It is assumed that this is the result of inappropriate intramolecular recombination of the unpaired X chromosome. Based on this, and a consideration of the structure of the deletion-inversion-deletion event, we propose a parsimonious mechanism (fig 2-D-G) whereby two crossovers form a serpentine intermediate which is then resolved to give two separate molecules. One is a linear product consisting of the deletion-inversion-deletion X chromosome inherited by the grandson, our patient. The other is a circular molecule comprising both deleted regions in an inverted relationship to one another (with respect to their wild type orientation), which would not be retained. The only comparable case of which we are aware is an instance of a slightly simpler deletion-inversion event involving the glycoprotein IIIa gene in a case of Glanzmann thrombocytopenia. Here the authors posit a related S shaped intermediate in their proposed mutational mechanism. A notable distinction between the haemophilia A inversions and the mutation described here is the degree and extent of similarity which drives the recombination; the former involves sequences exhibiting >99% identity over 9.5 kb, while the latter involves 80% identity over 130 bp. The mutational frequencies differ accordingly.

ACKNOWLEDGEMENTS
We are grateful to the Generation Trust and Guy's and St Thomas' Charitable Foundation for supporting this work, and would like to thank Dr G Hartnoll, Consultant Paediatrician, Lewisham University Hospital, Dr M Champion, Consultant in Paediatric Metabolic Medicine, Dr S Robb, Consultant Paediatric Neurologist, Guy's Hospital, and Dr C Buchanan, Consultant Paediatric Endocrinologist, King's College Hospital, who made the diagnosis of complex GKD and referred the family to us. Immunofluorescence analysis of dystrophin in muscle biopsy material was performed by Dr S Al-Sarraj at the Institute of Psychiatry, Denmark Hill, London. We are grateful to Dr Peter Green and Dr Elizabeth Green at Guy's Hospital for useful discussions regarding inversion mutagenesis.

Supplementary information can be found on the journal website (www.jmedgenet.com).