Molecular pathology of familial hypertrophic cardiomyopathy caused by mutations in the cardiac myosin binding protein C gene

B Yu, J A French, L Carrier, R W Jeremy, D R McTaggart, M R Nicholson, B Hambly, C Semsarian, D R Richmond, K Schwartz, R J Trent

Abstract
DNA studies in familial hypertrophic cardiomyopathy (FHC) have shown that it is caused by mutations in genes coding for proteins which make up the muscle sarcomere. The majority of mutations in the FHC genes result from missense changes, although one of the most recent genes to be identified (cardiac myosin binding protein C gene, MYBPC3) has predominantly DNA mutations which produce truncated proteins. Both dominant negative and haplonsufficiency models have been proposed to explain the molecular changes in FHC. This study describes two Australian families with FHC caused by different mutations in MYBPC3. The first produces a de novo Asn755Lys change in a cardiac specific domain of MYBPC3. The second is a Gln969X nonsense mutation which results in a truncated protein. Neither mutation has previously been found in the MYBPC3 gene. The consequences of DNA changes on the function of cardiac myosin binding protein C are discussed in relation to current molecular models for this disorder. (J Med Genet 1998;35:205–210)

Keywords: familial hypertrophic cardiomyopathy; MYBPC3 gene; protein

The autosomal dominant disorder familial hypertrophic cardiomyopathy (FHC) is characterised by excessive myocardial hypertrophy in the absence of abnormal haemodynamic load. In 1990, DNA studies first showed that FHC was caused by a mutation in a cardiac muscle gene.1 Subsequently, extensive molecular analyses have led to the proposal that FHC is a disease of the sarcomere.2 Mutations in several genes which code for sarcomeric proteins have now been described, including cardiac β myosin heavy chain on chromosome 14q11 in ~30% of FHC families, cardiac troponin T on chromosome 1q3 in ~15% of families, α tropomyosin on chromosome 15q2 in <3% of families, and cardiac myosin binding protein C (MYBPC3) on chromosome 11p11 in ~20% of families.1-7 A form of FHC which is partially coherited with the Wolff-Parkinson-White syndrome has been localised to chromosome 7q3.5 Mutations in myosin light chains (chromosomes 3p and 12q) may account for <1% of FHC families.5 The most recent sarcomere specific gene implicated in FHC is cardiac troponin I on chromosome 19p13.2-q13.2.10 With very few exceptions (one amino acid deletion and a donor site mutation in the troponin T gene, one amino acid deletion in troponin I), mutations observed in the cardiac β myosin heavy chain gene, the cardiac troponin T and I genes, and the α tropomyosin gene involve missense changes which affect conserved regions of each protein. The significance of these missense mutations in the pathogenesis of FHC has recently been reaffirmed by a transgenic mouse model, which carries an Arg403Gln mutation in the human cardiac β myosin heavy chain gene.11 Affected mice show both the histological features found in the human disorder and the complication of sudden death during exercise.

The pattern of mutations in MYBPC3 appears to be different from that for the other genes. Eight of the nine mutations reported in MYBPC3 to date result in a truncated protein.2-7 The present study describes two Australian families with novel mutations in the MYBPC3 gene. These are (1) a de novo missense change which does not produce a truncated protein and (2) a nonsense mutation leading to a premature stop codon. The position of the missense mutation in a highly conserved region of this gene suggests a mechanism by which the protein's function can be altered. The nonsense mutation is located towards the protein's carboxyl terminal end which is important for myosin and titin binding.

Methods
DNA HAPLOTYPES
In family F155, affected subjects were identified by conventional clinical criteria.12 Haplotypes were constructed for 15 subjects (12 family members and three spouses) and inferred for one additional person (III.1) who died, aged 47, from malignancy. Microsatellite markers on chromosomes 1, 14, and 15 were used to exclude known FHC loci (data not shown).13 Microsatellites for linkage studies at the 11p13-q13 region included D11S935, D11S905, D11S903, D11S1319, D11S1344, D11S1357, and D11S987.17 The lod score was calculated with the MLINK program, using the age dependent liability classes as reported previously.1 The chromosome 11 haplotypes were used to identify which allele cosegregated with the underlying MYBPC3 mutation, and, together with 10 microsatellites from the other...
three chromosomal loci, for confirmation of paternity. A similar investigative strategy was followed in family F237.

DNA MUTATION ANALYSIS

Family F155
MYBPC3 exons and their intronic splice sites were amplified by PCR using different sets of intronic primers and screened by SSCP. In the case of exon 23, a 281 bp product was obtained. The SSCP pattern was analysed on an 8-10% polyacrylamide gel and visualised by silver staining. The same product was directly sequenced using an ABI 373 automated sequencer (Perkin Elmer, Foster City). Allele specific oligonucleotides (ASOs) were designed for subsequent mutation detection in other family members as well as 150 normal controls. The normal and mutant ASO probes were 5'-AGTGAAGAACCCTGTGG-3' and 5'-GAAGAAACCTGTGGGCGG-3', respectively. Probe labelling, hybridisation, washing (at 47°C), and detection steps were carried out using an enhanced chemiluminescent system according to the manufacturer’s instructions (ECL 3' Oligolabelling and Detection System, Amersham Australia).

Family F237
SSCP for exon 27 gave a 274 bp PCR fragment, which was amplified by forward primer 5'-CAGTGGAGTGGGTTGTC-3' and reverse primer 5'-GTTGTAATGGGGGTCT-3'. This exon was sequenced. The technique of MS-PCR was applied to screen family members.15 In MS-PCR, the same forward

---

**Figure 1** Pedigree of family F155. Subjects with or without clinical evidence of FHC are shown as closed and open symbols respectively. Haplotypes are derived from seven microsatellite repeats from chromosome 11p13-q13. The MYBPC3 gene is located at 11p13.2 near D11S1344 and D11S1357. Haplotypes 7-6-2-4-4-3-6 and 7-6-2-4-4-3-6 apprehend all clinically affected FHC subjects and is also present in three clinically normal or equivocal subjects (II.3, III.3, and IV.3). All those with the 7-6-2-4-4-3-6 haplotype are further classified as Amin755Lys mutation positive (+) or negative (-). Deduction of the haplotype and mutation status in III.1 is indicated by parentheses. This subject died at the age of 47 from disseminated malignancy. Although there is no information about his cardiac status, he was an obligatory carrier.
Molecular pathology of familial hypertrophic cardiomyopathy

primer was used along with two fluorescent labelled reverse primers: wild type 5' (HEX)-GGCAAAAGGGCACCTACG-3' and mutant 5' (TET)-CTGgACAAAGGGCACCTACA-3' (g=deliberately introduced mismatch which releases the secondary structure of the primer). The mutant product was two bases longer than the wild type fragment and so was easily resolved using the ABI 310 PRISM Genetic Analyzer (Perkin Elmer, Foster City).

MYOSIN BINDING PROTEIN C ANALYSIS
In family F155, localisation of the mutated amino acid in the predicted three dimensional structure was obtained by comparing the conserved sequence patterns in the immunoglobulin I set (IgI) as described by Harpaz and Chothia. Partial amino acid sequences for both the cardiac and fast skeletal isoforms in human and chicken were obtained from GenBank. These sequences were aligned to maximise homology in the region of the mutation, that is, at motif 5 of human cardiac myosin binding protein C. Secondary structure alteration following a single amino acid residue replacement was predicted using the protein analysis function of the GCG program (Wisconsin) available through the Australian National Genomic Information Service.

Results

MOLECULAR

Family F155

Linkage analysis excluded the cardiac β myosin heavy chain, cardiac troponin T, and α troponymosin genes as causative in this family (data not shown). Because of the relatively small numbers available for study, linkage to the chromosome 11 FHC locus was only suggestive with a maximum two point lod score of 1.7 (at θ=0) for III.1 and his descendants (fig 1). Nevertheless, the result was sufficient to proceed to SSCP screening which showed an aberrant conformer in the PCR product for exon 23 of MYBPC3. DNA sequencing identified a transversion of C to A at nucleotide 2297 (GenBank X84075) which resulted in an asparagine to lysine substitution in amino acid residue 755 (Asn755Lys). ASO studies confirmed eight family members to be heterozygous for this mutation. The Asn755Lys mutation described was not found in 300 chromosomes from normal, unrelated controls.

Haplotype analysis showed that the 7-6-2-4-4-3-6 pattern cosegregated with the mutant MYBPC3 in III.1 and his descendants. However, the proband's paternal aunt (III.3) and great aunt (II.3) had the same haplotype but not the Asn755Lys change on ASO study (fig 1). This haplotype was derived using seven microsatellite markers distributed over a distance of ~24 cm on either side of MYBPC3. The result implies that a de novo mutation had occurred, but alternative explanations include non-paternity or that II.1 was homozygous for the haplotype. Non-paternity is unlikely since, based on the allele frequencies of the seven markers that make up the haplotype, the likelihood that it has occurred by chance is a conservative 1 in 105. An additional 10 highly polymorphic microsatellite markers from chromosomes 1, 14, and 15 failed to show any evidence of non-paternity among any of the people tested (data not shown). Homozygosity is similarly excluded since three distinct parental haplotypes (7-6-2-4-4-3-6, 4-7-1-3-3-4-8, and 5-1-3-7-3-4-5) were observed in generation III. The fourth haplotype (4-1-3-7-3-4-5) involves a recombination, but the finding of allele "4" at locus D11S935 excludes homozygosity, since the corresponding marker for the haplotype under discussion is "7".

Family F237

Linkage analysis excluded the cardiac β myosin heavy chain, cardiac troponin T, and α troponymosin genes as causative in this family (data not shown). No recombination was detected between loci D11S903 and D11S1357. The maximum two point lod score was +8.0 at 0=0. An aberrant conformer in the PCR product for exon 27 of MYBPC3 was identified through SSCP screening, and DNA sequencing identified a C to T transversion at nucleotide position 2837 (GenBank X84075). As a result, a premature stop codon (TAA) occurs in place of Gln (CAA) at amino acid position 969. Thirty members of family F237 showed this nonsense mutation (Gln969X) on MS-PCR.

PROTEIN STRUCTURE

Family F155

The predicted structure of cardiac myosin binding protein C is illustrated in fig 2A. Mutations described in this gene are shown in fig 2B and the regions of conservation in fig 2C (the location of the Asn755Lys mutation is indicated by an arrow). The missense mutation occurs in motif 5, which is not involved in myosin binding. This region is highly conserved between cardiac and fast skeletal proteins in both human and chicken muscle (fig 2C). Moreover, a 27 amino acid sequence present in motif 5 is only found in the cardiac specific isoform of myosin binding protein C. The effect of the Asn755Lys change on the protein’s conformation was examined and results are summarised in table 1. The GCG program indicates that a β turn occurs between Lys754 and Asp760. The Asn755Lys mutation interferes with this turn. The presence of a second positively charged lysine residue next to Lys754 would also be expected to cause further conformational change through charge repulsion. An alteration of hydrogen bonding capac-

| Table 1 Predicted changes in cardiac myosin binding protein C conformation owing to the Asn755Lys mutation |

<table>
<thead>
<tr>
<th>Residue No</th>
<th>753</th>
<th>754</th>
<th>755</th>
<th>756</th>
<th>757</th>
<th>758</th>
<th>759</th>
<th>760</th>
<th>761</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino acid*</td>
<td>V</td>
<td>K</td>
<td>N/K</td>
<td>P</td>
<td>V</td>
<td>G</td>
<td>E</td>
<td>D</td>
<td>Q</td>
</tr>
<tr>
<td>Charged</td>
<td>o</td>
<td>+</td>
<td>p/+</td>
<td>o</td>
<td>o</td>
<td>p</td>
<td>-</td>
<td>-</td>
<td>p</td>
</tr>
<tr>
<td>Conformation</td>
<td>Wild type</td>
<td>B</td>
<td>t</td>
<td>t</td>
<td>T</td>
<td>T</td>
<td>t</td>
<td>t</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Mutant</td>
<td>B</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

*V = valine; K = lysine; N = asparagine; P = proline; G = glycine; E = glutamic acid; D = aspartic acid;
Q = glutamine. Residue 755 is shown in wild (N) and mutant (K) types.
| t = neutral amino acid; + = positively charged amino acid; p = polar amino acid; − = negatively charged amino acid.
| B = amino acid residue predicted in β sheet; = amino acid residue predicted in turns and (T) amino acid residue predicted in specific β turn; = unpredictable.

Downloaded from http://jmg.bmj.com/ on July 6, 2017 - Published by group.bmj.com
A Human cardiac myosin binding protein C structure

<table>
<thead>
<tr>
<th>Normal</th>
<th>Cardiac specific region</th>
<th>Phosphorylation site</th>
<th>Linker</th>
<th>Myosin and titin binding site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutants</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B Mutants

1. Dup
2. M
3. T,S
4. T,D
5. M
6. T,S
7. T,D
8. T,S
9. T,S
10. T

C Comparison of partial amino acid sequences of motif 5 in different species and isoforms

Figure 2. (A) The proposed structure of the human cardiac myosin binding protein C. Four functional sites are indicated by solid lines. The extra IgI as the N-terminus is identified as motif 0 and the other seven IgI domains are numbered 1-5, 8, 10. Three fibronectin III domains are shown as squares 6, 7, and 9. (B) Nine mutations (mutants 1-9) previously reported are illustrated. The two novel mutations described in the present study are 10 and 11. The following symbols describe the mutations: T - truncated protein; S - splicing defect; Dup - duplication; M - missense mutation; * shows the position of missense changes, duplicated, or deleted regions. Thus, T,S would indicate a truncated protein arising from a splicing defect. Two alternatives are provided for mutations 5 and 7 because two mRNA species were found in each of these cases. Mutation 3 is the six amino acid duplication in motif 10 which is critical for myosin binding. (C) Partial amino acid sequences of motif 5 from both cardiac and fast skeletal MYBPC in human and chicken are listed, with their GenBank accession numbers in parentheses. The 755 amino acid residue in human MYBPC3 is indicated by an arrow.

Family F237

The Gln969X nonsense mutation results in a truncated protein with loss of motifs 8-10 (fig 2B).

Discussion

Myofibrils, which consist of repeating units known as sarcomeres, make up the contractile elements in muscle. A sarcomere comprises seven major proteins and several minor ones organized into thin and thick filaments. Muscle contraction and force generation results from the relative sliding between these filaments. Thin filaments consist predominantly of actin and lesser amounts of the regulatory proteins, tropomyosin, and troponin complex. Thick filaments are composed of myosin heavy chain, light chains, and other components including myosin binding protein C. The interaction of actin and myosin is initiated by calcium binding to the troponin complex. Actin binding stimulates myosin ATPase activity, resulting in a conformational change of the myosin molecule (power stroke) and force generation. Cardiac myosin binding protein C binds to myosin and, when phosphorylated, modulates the force of contraction.

The gene for human cardiac myosin binding protein C is MYBPC3. It encodes a protein of 1274 amino acids which belongs to the intracellular immunoglobulin (Ig) superfamily (GenBank X84075). The Ig superfamily is considerably divergent in terms of sequence and function, but all members share a common feature, which is a fold of ~100 amino acids assembled into two β sheets packed tightly to form a β sandwich. Based on similarities in sequence and structure of the ~100 amino acids, the Ig superfamily has been divided into four sets known as C1, C2, V, and I. Each of the two β sheets of the IgF set is made up of different strands. One β sheet has strands A, B, E, and D. The second β sheet has strands A', F, G, C, C'. The tightly packed structure of the β sandwich acts as a stable scaffold to facilitate highly specific protein-protein interactions. The proposed structure for the various isoforms (fast or slow skeletal, cardiac) of myosin binding protein C is similar and includes seven IgF domains and three fibronectin III domains (fig 2A). The cardiac isoform has an extra IgF domain (0) along with a cardiac specific region at the N-terminal end and a phosphorylation site between motifs 1 and 2. Myosin binding protein C has been localized to the C zone of the sarcomere. The highly conserved C-terminal end (motifs 7-10) of myosin...
Molecular pathology of familial hypertrophic cardiomyopathy

binding protein C has functional significance since it binds to myosin and titin. Evidence also suggests that myosin binding protein C can interact with both F actin and the motor domains of myosin and modulate the actin activated ATPase. Thus, myosin binding protein C plays a dual role in determining the sarcomere's structural integrity and in regulating myocardial contraction. The numerous missense changes described in MYBPC3 as well as the following sarcomere genes, cardiac β myosin heavy chain, cardiac troponins T, I, and tropomyosin, could perturb function through a dominant negative effect. A duplication of six amino acids at residues 1163-1168 is an exception to the protein truncating defects in MYBPC3. This mutation could interfere with myosin or troponin binding, disrupting the normal assembly of that protein in the sarcomere and has been hypothesised to cause dysfunction through a dominant negative mechanism. The present study describes a further example of a truncated protein caused by a mutation in MYBPC3. The Gln969X defect in family F237 leads to loss of motifs 8-10 at the protein’s carboxy end, which would interfere with binding by myosin or titin. As reported by Koshida et al., the truncated cardiac myosin binding protein C, which does not have its C-terminal 75 amino acids, is unable to bind to the myosin filaments in myogenic cell cultures. Lack of the full length protein affects the recruitment and organisation of newly synthesised myosin. Accumulation of disorganised myofibrils may then trigger the compensatory hypertrophic response.

In contrast, the Asn755Lys missense change in family F155 does not produce a truncated protein nor does it directly affect a region involved in myosin assembly in the sarcomere. Its potential role in disrupting cardiac myosin function can be determined by analysing its effect at the protein level. Amino acid residue 755 is located in a highly conserved region in motif 5 (fig 2C). It can be predicted to lie within a type I β turn that connects the F and G strands in one of the β sheets of the IgI domain.

Hydrogen bonding between the asparagine and the nearby Gly758 (fig 2C) is important in stabilising this β turn. Additional hydrogen bonding between Asn755 and the BC bend is predicted to help hold the two bends together. The one amino acid change from asparagine to lysine in the sequence could be expected to alter its structure, for example, by disrupting the hydrogen bond and decreasing the tightness of the turn through repulsion of the two lysine residues (table 1). In changing the conformation of motif 5, the Asn755Lys mutation could alter its specificity or capability to interact with other proteins.

The Asn755Lys change in family F155 has occurred de novo. Without DNA samples from III.1 and his parents, it is difficult to determine whether the mutation first occurred in II.1 or III.1 (fig 1). However, it is likely to have started in III.1 since a de novo change in II.1 would also require the presence of germinal mosaicism to explain the findings in the next generation. In itself, the de novo finding implies that it is pathogenic because a single base change between II.3, III.3, and other carriers of the 7-6-2-4-4-3-6 haplotype is sufficient to produce the FHC phenotype. The different locations for the 11 mutations described in MYBPC3 do not define a specific mutation hot spot. However, a cardiac specific IgI domain in MYBPC3 is the site for a disease causing mutation in family F155 and other mutations may be found in this region.

The present study shows how the function of a DNA mutation can be predicted by examining its effect on protein structure. Molecular characterisation of further mutations in the genes causing FHC is required to enable a better understanding of the phenotype/genotype correlations in this disorder.

This work was supported by grants from the NHMRC of Australia, the Rotary Australia Health Research Fund, INSERM, and the Association Française Contre les Myopathies. We are grateful to a recipient of a Dora Lush (Biomedical) Postgraduate Scholarship from the NHMRC. We thank Ms Corinne Crusaud for the sequence analysis. This work would not have been possible without the help and cooperation of families F237 and F155.


15 Harpaz Y, Chothia C. Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. J Mol Biol 1994;238:525-39.


Molecular pathology of familial hypertrophic cardiomyopathy caused by mutations in the cardiac myosin binding protein C gene.

B Yu, J A French, L Carrier, R W Jeremy, D R McTaggart, M R Nicholson, B Hambly, C Semsarian, D R Richmond, K Schwartz and R J Trent

doi: 10.1136/jmg.35.3.205

These include:

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/