Gamma-D crystallin gene (CRYGD) mutation causes autosomal dominant congenital cerulean cataracts

E Nandrot, C Slingsby, A Basak, M Cherif-Chefchaouni, B Benazzouz, Y Hajaji, S Boutayeb, O Griboval, L Arbogast, A Berraho, M Abitbol*, L Hilal*

ORIGINAL ARTICLE

See end of article for authors’ affiliations

Correspondence to:
Dr M Abitbol, Centre de Recherches Thérapeutiques en Ophtalmologie, Faculté de Médecine Necker-Enfant Malades, 156 Rue de Vaugirard, 75015 Paris cedex, France; abitbol@necker.fr or Pr L Hilal, Laboratoire de Génétique et Biologie Moléculaire, Département de Biologie, Faculté des Sciences, Université Ibn Tofail, Kénitra, Morocco; lhilal@yahoo.fr

19 December 2002

Accepted for publication 13 January 2003

C
ongenital cataracts (ADCC) are a significant cause of bilateral visual impairment in childhood. We mapped the gene responsible for autosomal congenital cerulean cataracts to chromosome 2q33-35 in a four generation family of Moroccan descent. The maximum lod score (7.19 at recombination fraction θ=0) was obtained for marker D2S2208 near the γ-crystallin gene (CRYG) cluster. Sequencing of the coding regions of the CRYGA, B, C, and D genes showed the presence of a heterozygous C>A transition in exon 2 of CRYGD that is associated with cataracts in this family. This mutation resulted in a proline to threonine substitution at amino acid 23 of the protein in the first of the four Greek key motifs that characterise this protein. We show that although the x-ray crystallography modelling does not indicate any change of the backbone conformation, the mutation affects a region of the Greek key motif that is important for determining the topology of this protein fold. Our data suggest strongly that the proline to threonine substitution may alter the protein folding or decrease the thermodynamic stability or solubility of the protein. Furthermore, this is the first report of a mutation in this gene resulting in autosomal dominant congenital cerulean cataracts.

MATERIALS AND METHODS

Family data
We have studied a large Moroccan family with ADCCC. The family consists of 19 affected and 24 unaffected subjects spanning four generations. Appropriate informed consent was obtained from all patients participating in the study. Clinical and ophthalmological examination of patients affected by cerulean cataract as well as unaffected family members were carried out as described previously. The diagnosis of cerulean cataract was confirmed in each affected patient by four independent ophthalmologists. There was no history of other ocular or systemic abnormalities in this family. The clinical aspects of the present study have been published recently.

*The last two authors contributed equally to this work
Blood samples were collected from 43 family members. Genomic DNA was extracted by standard techniques.

**Genotyping and linkage analysis**

The genotyping was performed as described previously using firstly 30 microsatellite markers, corresponding to 13 known candidate loci for autosomal dominant congenital cataracts, and then 21 markers localised to 2q33-35. The oligonucleotide primer sequences were taken from Généthon (http://www.genethon.fr). Two point disease to marker linkage analysis was conducted by MLINK from the FASTLINK (version 3.0P) software package. The mode of inheritance was considered to be autosomal dominant with full penetrance. The gene frequency was set at 1/20 000. As the allelic frequencies of the polymorphic markers were unknown in the Moroccan population, they were considered to be equal. Order and genetic distances were taken from the Marshfield database (http://research.marshfieldclinic.org) and ensembl genome data resources (http://www.ensembl.org). Multipoint analysis was computed using Genehunter software.

**DNA sequencing**

Genomic DNA samples from all affected and unaffected family members of the pedigree and from control subjects were screened for mutations in CRYGA, B, C, and D by direct cycle sequencing. Gene specific PCR primers were used to amplify the three exons and flanking introns sequences of CRYGA, CRYGB (sequences available upon request), CRYGC, and CRYGD. PCR products were purified by means of “invitro” rapid PCR purification systems and were sequenced on ABIprism A310 and ABI A377 automated sequencer (PE Biosystems, USA) using the original and additional internal primers.

**Molecular modelling**

A model of the P23T mutant structure was built based on the human γD crystal coordinates using the program “O” by placing the mutant side chain in the most favourable conformation using the side chain database within the program and using the lego-side chain option. The model was energy minimised using 200 cycles of Powell energy minimisation using the CNS program and the quality of the model was checked with Procheck.

**RESULTS**

We analysed, clinically and genetically, a large Moroccan family affected by ADCCC. The congenital cataract phenotype was clearly distinct from the phenotypes of the lamellar, coralliform, aculeiform, Coppock-like, and polymorphic congenital cataracts phenotypes.

Forty-three members of the pedigree, including 20 affected subjects, 17 unaffected family members, and six unaffected spouses (fig 1A), were genotyped. The preliminary linkage analysis performed in this family using 30 microsatellite markers allowed us to exclude 12 candidate loci for ADCC (1pter-p36.13, 1q21.2, 1q21.2, 3p21.2-22.3, 10q23.3-25, 1q21-14.1, 13q11-12, 16q22.1, 17p12-13, 17q11-q21.2.17q24, 21q22.3, and 22q11.2-q12.2) (data not shown). Positive two point lod scores were obtained for markers D2S72 (Zmax=3.47 at θ=0.01) and CRYGA (Zmax=1.75 at θ=0.00) (table 1). D2S72 and CRYGA flank the 4.5 cm 2q33-35 interval, where the loci corresponding to polymorphic congenital cataracts (MIM 601286), Coppock-like cataracts (MIM 123660, MIM 116200), juvenile onset punctate cataracts (MIM 123690), aculeiform congenital cataracts (MIM 115700), and variable zonular pulverulent cataracts have been previously mapped. Twenty-one additional markers spanning this critical interval were subsequently used for further genotyping of all family members. The two point lod scores for 10 of these 21 markers are summarised in table 1. Significant positive lod scores (Zmax>3 at θ=0) were found for 11 markers. The maximum two point lod score was obtained with marker D2S2208 (Zmax=7.19 at θ=0) (table 1). Multipoint analysis with the most informative markers confirmed that this locus mapped to chromosome 2q33-q35 (data not shown).

Haplotype analysis showed that the affected patients of this family shared a common haplotype involving 12 markers (D2S2237, D2S155, D2S2358, D2S355, D2S2192, D2S325, D2S242, D2S2208, D2S157, CRYGA, D2S2322, D2S128). Critical recombination events were detected in affected subjects III.6, III.7, III.11, IV.21, and IV.5. This allowed us to define a disease gene containing an interval of about 11.5 cM between markers D2S72 and D2S3261 (fig 1B). All affected subjects had an affected parent, and no unaffected subjects carried the disease haplotype. Thus, penetrance appears to be virtually complete in this family. Several candidate genes have been mapped in this interval, the obvious one being the γ-crystallin gene cluster, CRYG, which has been mapped to 2q33-35. Another crystallin gene, CRYBA2, has been mapped to the 2q34-36 region. CRYBA2 is 10.5 cm telomeric to CRYGA and 1 cM telomeric to D2S173. ELOX3, a major developmental gene, is localised 1 cM telomeric to D2S126 and is thus outside the critical interval. The observation of recombination events, which are centromeric to these two genes, in our family excluded the CRYBA2 and ELOX3 genes as candidate genes for the cataract phenotype associated with the cerulean blue dot cataract type. This type differs from those previously described and is characterised by early onset (diagnosed at birth) and by a faster progression.

| Table 1: Two point lod scores (Z) for linkage between autosomal dominant congenital cataract CCA3 and markers in the chromosomal region 2q33-35 |
|---|---|---|---|---|---|---|---|---|---|---|
| Marker | Genetic distance* (Mb) | Lod score at θ |
| | | 0.0 | 0.01 | 0.05 | 0.1 | 0.2 | 0.3 | 0.4 |
| D2S116 | 3.09 | --- | 2.22 | 3.22 | 3.27 | 2.70 | 1.77 | 0.70 |
| D2S72 | 0.87 | --- | 3.47 | 3.85 | 3.72 | 3.07 | 2.17 | 1.08 |
| D2S2237 | 1.32 | 4.61 | 4.54 | 4.24 | 3.85 | 2.99 | 2.02 | 0.93 |
| D2S155 | 1.16 | 5.87 | 5.77 | 5.37 | 4.84 | 3.72 | 2.50 | 1.18 |
| D2S2192 | 0.61 | 5.46 | 5.37 | 5.01 | 4.53 | 3.51 | 2.38 | 1.14 |
| D2S2208 | 0.36 | 7.19 | 7.09 | 6.67 | 6.11 | 4.86 | 3.44 | 1.82 |
| CRYGA | 5.95 | 1.75 | 1.71 | 1.54 | 1.33 | 0.88 | 0.46 | 0.12 |
| D2S128 | 1.23 | 5.72 | 5.64 | 5.30 | 4.84 | 3.84 | 2.70 | 1.41 |
| D2S2256 | 2.31 | --- | 4.82 | 5.07 | 4.77 | 3.79 | 2.53 | 1.09 |
| D2S173 | 3.16 | --- | 1.29 | 1.78 | 1.82 | 1.56 | 1.12 | 0.59 |
| D2S126 | --- | 1.18 | 2.84 | 3.18 | 2.85 | 2.03 | 0.95 |

*Between the adjacent mapped markers.
Given our linkage analysis results and the fact that γ-crystallin genes are expressed early in development and are associated with hereditary cataracts in mice and humans, we focused on these genes. The analysis of the sequences of all three exons of CRYGA, CRYGB, CRYGC, and CRYGD did not show any disease causing mutations in the CRYGA, CRYGB, or CRYGC gene in any of the members of this large family. Only the single nucleotide polymorphisms previously described were identified, like 2437C/T (M19364) (P64P) in exon 3 of CRYGB, 5391C>A (M193364) (L111I) in exon 3 of CRYGB, 18229A>G (M19364) (IVS-70 intron A) in CRYGC, and 2867T>C (K03005) (Y16Y) in exon 2 of CRYGD. A unique heterozygous C>A transversion was identified at nucleotide 305 in exon 2 (K03005) of the CRYGD gene in all affected family members. This transversion led to the replacement of a proline residue at amino acid 23 by a threonine (P23T) in the first “Greek key motif” (motif 1) of the γD-crystallin protein. This substitution was not found in any crystallin, whatever the species analysed, at this position. Proline has a hydrophobic side chain often associated with turns in polypeptide chain conformation, whereas the threonine side chain has both hydrophilic and hydrophobic functions. Furthermore, a threonine is never found in any crystallin, whatever the species analysed, at this position. Proline is known to have a strong ability to break β-strands in soluble proteins. These data strongly suggest that the 305C>A transversion is indeed the CCA3 causing mutation rather than a rare benign polymorphism. It is noteworthy that the same mutation has recently been suggested to cause an autosomal dominant congenital form of lamellar cataracts in a proband and her affected father in a small nuclear Indian family. The difference between the phenotype reported in this Indian family and that observed in our Moroccan family may be related to the effect of an unknown modifier gene or to sequence variations within regulatory regions that could

**FIGURE 1** (A) Pedigree of the large Moroccan family affected by autosomal dominant congenital cerulean cataracts (CCA3). (B) Markers of the chromosomal region 2q33-q35 used for linkage analysis. a Common haplotype present in all affected family members. b Haplotypes observed in patients III.6, III.7, III.11, IV.5 and IV.21, displaying critical recombination events. Asterisks indicate subjects who underwent linkage analysis.

**DISCUSSION**

In this report, after excluding most known loci corresponding to ADCC, we identified a third locus (CCA3) on 2q33-35, associated with the cerulean blue dot cataract type in a large Moroccan family. We then found a C>A transversion in exon 2 of CRYGD only in all affected members of the large family studied. Crystallin genes encode a superfamily of major soluble structural proteins in the lens. There are three major classes of crystallins in humans, the α-, β-, and γ-crystallins. All three types of crystallins are β pleated sheets. Protein sequence analysis has shown homology between the β- and γ-crystallins. The γ-crystallin gene cluster comprises six genes; γA, γB, γC, γD, γE, and γF. Different levels of expression of these genes are associated with the lens refractive index characteristics of the vertebrate enabling adaptation to varying optical requirements. In mammals, these genes each consist of three exons. Only γC- and γD encode abundant lens γ-crystallins in humans. γE and γF are pseudogenes with in frame stop codons. γD is one of the only two γ-crystallins to be expressed at high concentrations in the fibre cells of the embryonic human lens. These cells subsequently form the lens nucleus fibres.

The identified C>A transversion found exclusively in all affected members of the pedigree results in a P23T substitution. Proline has a hydrophobic side chain often associated with turns in polypeptide chain conformation, whereas the threonine side chain has both hydrophilic and hydrophobic functions. Furthermore, a threonine is never found in any crystallin, whatever the species analysed, at this position. Proline is known to have a strong ability to break β-strands in soluble proteins. These data strongly suggest that the 305C>A transversion is indeed the CCA3 causing mutation rather than a rare benign polymorphism. It is noteworthy that the same mutation has recently been suggested to cause an autosomal dominant congenital form of lamellar cataracts in a proband and her affected father in a small nuclear Indian family. The difference between the phenotype reported in this Indian family and that observed in our Moroccan family may be related to the effect of an unknown modifier gene or to sequence variations within regulatory regions that could
affect the expression of the γD-crystallin gene. This phenotypic heterogeneity caused by an identical genetic crystallin mutation is strongly reminiscent of that previously reported for the βB2-crystallin gene. The same chain termination mutation in the βB2-crystallin gene caused autosomal dominant congenital Coppock-like cataracts in a four generation Swiss family and autosomal dominant congenital cerulean cataracts (CCA2) in a very large American family. These data further strengthen the hypothesis that the γD-crystallin P23T mutation is pathogenic and prompted us to try to obtain further insights into the molecular mechanisms underlying cataractogenesis in the large Moroccan kindred studied.

**Figure 2** Mutation analysis of CRYGD. (A) Sequence chromatograms of the wild type allele showing a proline (CCC) at amino acid 23. (B) Sequence chromatograms of the mutant allele showing the C to A transversion that changed proline 23 to threonine. (C) Sequence alignment of the γ-crystallin proteins in different species. Shown below are the amino acid residues 12-33 and 45-53 of human γD-crystallin. Proline 23 of human CRYGD protein is underlined and appears in bold type. The arrow indicates the P23T mutated position. Residues in the other γ-crystallins identical to those at homologous positions of human CRYGD are indicated by dots.

**Figure 3** (A) Cartoon topology diagram of the characteristic domain structure found in all γ-crystallins showing how two “Greek key motifs”, each comprising four β-strands (a, b, c, and d), associate to form two β-sheets. The red circle shows the site of the P23T mutation in the N-terminal domain of γD-crystallin. (B) Schematic view of the protein fold determined by x-ray crystallography centred on residue P23 and looking along β-sheet 2 of the normal CRYGD protein. (C) The same view centred on the mutation site but based on a predicted structure modelled on the crystallographic coordinates of the normal CRYGD protein. Light blue = β-sheet strands, green = loops connecting β-sheet strands, dashed green lines = hydrogen bonds between the region adjacent to the β-sheet c strand and residues in the folded ab β-hairpin, dashed yellow lines = local hydrogen bond in the region of the β-sheet c strand, red = oxygen atoms and blue = nitrogen atoms.
The 3D structures of several members of the monomeric γ-crystallin family are known, showing that the polypeptide folds into four similar Greek key motifs. These motifs then form two similar domains joined by a linker (recently reviewed in Jaenicke et al40). The 3D structure of human γD-crystallin has now been determined at a very high resolution (1.25 Å) by x ray crystallography. We may try to gain some insight into the pathogenicity of the P23T γD-crystallin mutation from the crystallographic structure of the native human protein. Its conformation is very similar to that of γ-crystallins from other species and, as expected, very similar to that of the orthologous bovine γD. Each γ-crystallin domain is formed from two consecutive Greek key motifs and each motif comprises four consecutive β-strands. Thus, the N-terminal domain can be described as comprising \( \alpha_1 \beta_1 \alpha_2 \\beta_2 \alpha_2 \beta_2 \alpha_2 \beta_2 \beta_1 \) and \( \alpha_1 \beta_1 \alpha_2 \\beta_2 \alpha_2 \beta_2 \alpha_2 \beta_2 \beta_1 \) β-strands. These eight β-strands form two β-sheets (\( \beta_1 \alpha_1 \beta_1 \alpha_2 \) and \( \beta_2 \alpha_2 \beta_2 \alpha_2 \beta_2 \beta_1 \) \( \beta_1 \alpha_1 \beta_1 \alpha_2 \) and \( \beta_2 \alpha_2 \beta_2 \alpha_2 \beta_2 \beta_1 \)) that pack together to form a β-sandwich domain. Although β-sandwich domains are common in proteins, in β-crystallins they are characterised by their high internal conformational symmetry and by a conserved folded hairpin structure between \( \alpha_1 \) and \( \beta_1 \) strands.

The mutation affects a region of the first Greek key motif that is located immediately after the \( \alpha_1 \) strand, a region that is important for determining the topology of this motif. If the native fold after crossing from one β-sheet to the next the polypeptide chain has to get back to the original β-sheet to place the \( \beta_1 \) strand within it. Obviously, a single Greek key cannot fold on its own; it is dependent on the concomitant folding of its partner motif. The degree of sequence conservation of residue 23 in motif 1 and the corresponding region in the other three motifs gives an idea of the importance of this region. Although not absolutely conserved, this residue tends to be a proline or a serine in motifs 1, 2, and 3, whereas it is an arginine in motif 4.

The mutation site thus substitutes a threonine residue into a serine in motifs 1, 2, and 3, whereas it is an arginine in motif 4. The degree of sequence conservation of residue 23 in motif 1 and the corresponding region in the other motifs gives an idea of the importance of this region. Although not absolutely conserved, this residue tends to be a proline or a serine in motifs 1, 2, and 3, whereas it is an arginine in motif 4.

The mutation site thus substitutes a threonine residue into a serine in motifs 1, 2, and 3, whereas it is an arginine in motif 4. The degree of sequence conservation of residue 23 in motif 1 and the corresponding region in the other three motifs gives an idea of the importance of this region. Although not absolutely conserved, this residue tends to be a proline or a serine in motifs 1, 2, and 3, whereas it is an arginine in motif 4.
Evidence of clinical and genetic heterogeneity in autosomal dominant juvenile-onset, progressive cataract locus on chromosome 3q21-q22 is associated with two families with dominantly inherited cataracts.

Mutations in MIP underlie autosomal dominant ‘polymorphic’ and autosomal dominant “zonular pulverulent” cataract, on chromosome 1q.

A nonsense mutation in the human connexin50 gene (GJA8) underlies congenital cataract.

The x-ray structure of a beta-crystallin (1.25Å) and the R58H mutant (1.15Å) associated with a chain termination mutation in the human beta-crystallin crystallization. A missense mutation in the human connexin46 gammaC-crystallin gene is associated with autosomal dominant variable lamellar and Marner cataract.

Associated with a chain termination mutation in the beta-crystallin region on chromosome 22. Maps to the beta crystallin region on chromosome 22.


Linkage computations.

A 5-base insertion in the human beta-crystallin gammaC-crystallin T5P mutant. A 5-base insertion in the human connexin50 gene (GJA8) underlies congenital cataract.

Gamma-D crystallin gene (CRYGD) mutation causes autosomal dominant congenital cataracts

E Nandrot, C Slingsby, A Basak, M Cherif-Chefchaouni, B Benazzouz, Y Hajaji, S Boutayeb, O Gribouval, L Arbogast, A Berraho, M Abitbol and L Hilal

doi: 10.1136/jmg.40.4.262

Updated information and services can be found at:
http://jmg.bmj.com/content/40/4/262

These include:

**References**
This article cites 56 articles, 15 of which you can access for free at:
http://jmg.bmj.com/content/40/4/262#BIBL

**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.

**Topic Collections**
Articles on similar topics can be found in the following collections

- Eye Diseases (298)
- Genetic screening / counselling (886)
- Molecular genetics (1254)
- Clinical genetics (256)

**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/