ORIGINAL ARTICLE

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

E Nandrot, C Slingsby, A Basak, M Cherif-Chefchaouni, B Benazzouz, Y Haijji, S Boutayeb, O Gribouval, L Arbagost, A Berraho, M Abitbol*, L Hilal*

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; hlihal@yahoo.fr

Revised version received 19 December 2002
Accepted for publication 13 January 2003

Congenital cataracts are a major 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 CRYG, B, C, and D genes showed the presence of a heterozygous C>A transition in exon 2 of CRYG 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.

C

ongenital cataracts (ADCC) are a significant cause of visual impairment in childhood. The estimated prevalence is 1-6/10 000 live births. In non-consanguineous populations, most cases of inherited non-syndromic cataracts show an autosomal dominant (AD) mode of transmission, but X linked and autosomal recessive forms have also been observed. Inherited cataracts are clinically highly heterogeneous and show considerable inter- and intrafamilial variability. To date, more than 15 independent loci and genes on different chromosomes have been shown to be associated with autosomal dominant congenital cataracts (ADCC), and causative mutations have been identified in several distinct genes, including genes encoding crystallins, lens specific connexins (gap junction proteins), aquaporin (MIP), and cytoskeletal structural proteins, as well as in the homeobox gene PITX3 and the heat shock transcription factor 4 gene (HSF4). Congenital cataracts are genetically heterogeneous as clinically identical cataracts have been mapped to different loci. Conversely, an identical mutation can result in different phenotypes.

Cerulean cataracts, which were first described by Vogt in 1922, are characterised by early onset, progressive development of bilateral lens opacities and complete penetrance. Affected newborns appear to be asymptomatic until the age of 18 to 24 months. At this age they can be clinically diagnosed by slit lamp examination, which shows the presence of tiny blue or white opacities that first develop in the superficial layers of the fetal lens nucleus. Armitage et al suggested that cerulean cataracts should be classified as developmental cataracts rather than as congenital cataracts. The opacities spread throughout the adult lens nucleus and cortex and form concentric layers with radially orientated central lesions.

Two cerulean ADCCs (ADCCC) have been mapped in distinct, large families: one to 17q24 (CCA1, MIM 115660) and the other to 22q11.2-q12.2 (CCA2, MIM 601547), which is close to the β-crystallin gene (CRYBB2, MIM 123620). Litt et al identified a chain termination mutation in the CRYBB2 gene and showed that this mutation is associated with the CAA2 type. However, no mutated genes have yet been found to be associated with the CCA1 type. We recently reported a large Moroccan pedigree with an unusual form of cerulean blue dot cataract, characterised by a very early onset of lens opacities and a rapid progression towards blindness in the absence of rapid surgical intervention. In this family, we excluded any linkage to previously mapped loci for ADCCC (17q24 and 22q). We report here the linkage of the ADCCC affecting this Moroccan family to a region of chromosome 2q33-q35, spanning the γ-crystallin gene cluster. We also identified a missense mutation in the CRYGD gene, causing this precocious and rapidly progressing form of ADCCC.

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.
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.geneithon.fr). Two point disease to marker linkage analysis was performed 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 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 Genethunter 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, CRYGC, and D by direct cycle sequencing. Gene-specific PCR primers were used to amplify the three exons and flanking introns sequences of CRYGA, CRYGB, and CRYGD (sequences available upon request). CRYGC, and CRYGD. PCR products were purified by means of ‘in vitro-gen’ 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, corifolliform, 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, 1q21.3, 2p12.1-22.3, 10q23.3-25, 1q21.1-1.1, 13q11-12, 16q22.1, 17p12-13, 17q12-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), 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>3.0 at θ=0) were found for 11 markers. The maximum two point lod score was obtained with marker D2S2208 (Zmax=7.19 at θ=0.00) (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 D2S2361 (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. BAX3, 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 BAX3 genes as candidate genes for the cataract phenotype affecting this family. Thus, this linkage analysis identified a third locus (CCA3) associated with the congenital 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**

<table>
<thead>
<tr>
<th>Marker</th>
<th>Genetic distance (Mb)</th>
<th>Lod score at θ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0</td>
<td>0.01</td>
</tr>
<tr>
<td>D2S116</td>
<td>3.09</td>
<td>2.24</td>
</tr>
<tr>
<td>D2S72</td>
<td>0.87</td>
<td>3.47</td>
</tr>
<tr>
<td>D2S2237</td>
<td>1.32</td>
<td>4.61</td>
</tr>
<tr>
<td>D2S155</td>
<td>1.16</td>
<td>5.87</td>
</tr>
<tr>
<td>D2S2192</td>
<td>0.61</td>
<td>5.46</td>
</tr>
<tr>
<td>D2S2208</td>
<td>0.36</td>
<td>7.19</td>
</tr>
<tr>
<td>CRYGA</td>
<td>5.95</td>
<td>1.75</td>
</tr>
<tr>
<td>D2S128</td>
<td>1.23</td>
<td>5.72</td>
</tr>
<tr>
<td>D2S2208</td>
<td>2.31</td>
<td>4.82</td>
</tr>
<tr>
<td>D2S173</td>
<td>3.16</td>
<td>1.29</td>
</tr>
<tr>
<td>D2S126</td>
<td>1.18</td>
<td>2.84</td>
</tr>
</tbody>
</table>

*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 2 of CRYGB, 5391C>A (M193364) (L111I) in exon 3 of CRYGB, 18229A>G (M19364) (IVS-70 intron A) in CRYGC, and 286T>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 of the unaffected members of the pedigree. This mutation cosegregated strictly with the trait in the kindred studied. This demonstrates that the identified C>A transversion found exclusively in all affected members of the pedigree results in a P23T substitution. The threonine residue at position 23 of CRYGD protein is conserved across species, suggesting a potential role in the biological function of the gene. The identification of this mutation in the large Moroccan family provides a valuable resource for further studies on the genetic basis of hereditary cataracts.

**DISCUSSION**

In this report, after excluding most known loci corresponding to ADCC, we identified a third locus (CCA3) on 2q33-35, associated with the catarue 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. Crystalin 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. The amino acid change at position 23 is a non-conservative amino acid change as proline has a hydrophobic side chain.
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\textsuperscript{14} and autosomal dominant congenital cerulean cataracts (CCA2) in a very large American family.\textsuperscript{10,32,57}

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 βγ-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 al.6). The 3D structure of human γD-crystallin has now been determined at a very high resolution (1.25 Å) by x ray crystallography.7 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 species8 and, as expected, very similar to that of the orthologous bovine γD.9 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 a1 b1 c1 d1 and a2 b2 c2 d2 β-strands. These eight β-strands form two β-sheets (b1 a1 d1 c2 and b2 a2 d2 c1) that pack together to form a β-sandwich domain. Although β-sandwich domains are common in proteins, in βγ-crystallins they are characterised by their high internal conformational symmetry and by a conserved folded hairpin structure between a and b β-strands. The mutation affects a region of the first Greek key motif that is located immediately after the c1 strand, a region that is important for determining the topology of the native domain. After crossing from one β-sheet to the next the polypeptide chain has to get back to the original β-sheet to place the d1 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 position in terms of the fold topology, a site that has been occupied, however, by a very similar residue, serine. Fig 3 shows the N-terminal domain along the second β-sheet from the edge of the short c1 strand and centred around residue 23. In wild type human γD-crystallin, the carbonyl oxygen of Pro 23 forms a hydrogen bond with the backbone amide nitrogen of Asn 49 (fig 3B). This backbone interaction helps to stabilise the folded abc hairpin onto the sheet and is thus an important tertiary structure determinant. However, as it involves only backbone interactions, it is not affected by side chain alterations, assuming that in the mutant structure the backbone is in the same conformation as in the native structure. When threonine is modelled into this position, the new side chain is readily accommodated into the space without changing the backbone conformation (fig 3C). Threonine is more polar than proline as it has a polar backbone amide side chain has to get back to the original β-sheet to place the d1 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 position in terms of the fold topology, a site that has been occupied, however, by a very similar residue, serine. Fig 3 shows the N-terminal domain along the second β-sheet from the edge of the short c1 strand and centred around residue 23. In wild type human γD-crystallin, the carbonyl oxygen of Pro 23 forms a hydrogen bond with the backbone amide nitrogen of Asn 49 (fig 3B). This backbone interaction helps to stabilise the folded abc hairpin onto the sheet and is thus an important tertiary structure determinant. However, as it involves only backbone interactions, it is not affected by side chain alterations, assuming that in the mutant structure the backbone is in the same conformation as in the native structure. When threonine is modelled into this position, the new side chain is readily accommodated into the space without changing the backbone conformation (fig 3C). Threonine is more polar than proline as it has a polar backbone amide nitrogen and a side chain hydroxyl function. In this predicted model the mutant side chain appears to strengthen the interaction with tyrosine 50. The aromatic side chain of this residue can play an important role in stabilising the folded abc hairpin. Threonine 23 can make a stronger hydrogen bond with the side chain hydroxyl function of Tyr 50 using its backbone amide as well as an additional hydrogen bond with its own side chain hydroxyl. It is difficult to imagine that the mutation has such a catastrophic effect on the 3D structure of the human γD-crystallin based on this model. There are three possible scenarios: the protein does not fold properly because a side chain with a branched Cβ atom has a propensity to extend the β-strand rather than make the necessary turn; the substitution of threonine decreases the thermodynamic stability of the N-terminal domain in a subtle way so that the protein is more prone to denaturation; or thirdly the threonine side chain somehow decreases the solubility of the protein in water. The discovery of this mutation site, with its major clinical consequences, challenges our understanding of crystallin structure. This is clearly shown also by studies on other cataractogenic mutations of human γD-crystallin; recent biophysical studies suggest that the R14C γD-crystallin mutation increases the phase separation temperature and that the two γD-crystallin mutant proteins, R35H and R36S, crystallise far more readily than the native protein.12 However, none of these mutant proteins has been shown to change the conformation of the protein significantly. In contrast, the TSP CRYGC mutant protein induces obvious folding changes and decreases conformational stability.13 Further biophysical characterisation of the P23T mutant protein is important to provide more information about the nature of the disease and improve our understanding of the molecular photon of lens crystallins.

ACKNOWLEDGEMENTS

We are grateful to the members of the family for their participation in this study; we warmly thank Dr Jayanti Pande for critical reading of the manuscript. We thank Université René Descartes and Professor Patrick Berche, Dean of Faculté de Médecine Necker for his permanent support for the CERTO Laboratory. We thank Ibn Tofail University, Faculty of Sciences of Kénitra and Centre National pour la Recherche Scientifique et Technique de Morocco. We thank Professor Jean-Louis Dueler for his constant encouragement and efforts for the development of the CERTO Laboratory. We also thank Dr Cécile Mar- sac and Dr Maurice Menasche for their scientific support. We thank all members of the medical staff of the Department of Ophthalmology B (Hôpital des Spécialités de Rabat, Morocco). This work was mainly supported by a grant to the CERTO Laboratory from RETINA France also by a “Programme d’Appui à la Recherche Scientifique” (Pars Médicine No 151), by a research grant from INSERM (Programme PROGRES), and by several grants from Faculté de Médecine Necker and Université René Descartes. Emeline Nandrot has been supported by a PhD grant from RETINA-France for three years and by a PhD grant for one year from Fondation pour la Recherche Médicale. We gratefully acknowledge the financial support of the Medical Research Council, London.

Authors’ affiliations

E Nandrot, L Arbobast, M Abibost, Centre de Recherches Therapeutiques en Ophtalmologie de la Faculté de Médecine Necker, EA No 2502 du Ministère de la Recherche et de l’Enseignement Supérieur, Université René Descartes, Paris, France

C Slingsby, A Basak, Birbeck College, Department of Crystallography, Molé East Street, London WC1E 7HX, UK

M Cherif-Chefchaouni, A Berraha Département d’Oto-Neuro-Ophthalmologie, Service d’Ophtalmologie B, Hôpital des Spécialités, Rabat, Morocco

B Benazzouz, Y Hajaji, S Boutayeb, L Hilal, Département de Génétique et Biologie Moléculaire de la Faculté des Sciences, Université Ibn Tofail, Kénitra, Morocco

O Gribouval, Unité INSERM 423, CHU Necker-Enfants Malades, Paris, France

REFERENCES


10. Litt M, Carrero-Velanzuela R, LaMonticella DM, Schultz DW, Mitchell TN, Kramer P, Maumenee IE. Autosomal dominant cataract (located at 10q11.21-22) is
Juvenile-onset punctate cataracts caused by mutation of the gammaC-crystallin gene


Juvenile-onset punctate cataracts caused by mutation of the gammaC-crystallin gene