JBC RESEARCH ARTICLE

Cataract-causing Y204X mutation of crystallin protein CRYβB1 promotes its C-terminal degradation and higher-order oligomerization

Received for publication, May 11, 2023, and in revised form, June 16, 2023 Published, Papers in Press, June 24, 2023,
https://doi.org/10.1016/j.jbc.2023.104953

Xuping Jing[1,2,3,‡], Mingwei Zhu[1,2,‡], Xiaoyun Lu[1,2,‡], Ping Wei[1,2], Lingyu Shi[4], Bu-Yu Zhang[3,5], Yi Xu[1,2], Ya-Ping Tang[1,2,6,7,*], Dao-Man Xiang[4,*], and Peng Gong[3,*]

From the [1]Joint Laboratory for Translational Precision Medicine, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, Guangdong, China; [2]Joint Laboratory for Translational Precision Medicine, Wuhan Institute of Virology, and; [3]Key Laboratory of Special Pathogens and Biosafety, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan, Hubei, China; [4]Department of Ophthalmology, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, Guangdong, China; [5]University of Chinese Academy of Sciences, Beijing, China; [6]Goungdong Key Laboratory of Structural Birth Defects, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, Guangdong, China; [7]Department of Imaging, Affiliated Hospital 3, Zhengzhou University, Zhengzhou, Henan, China

Reviewed by members of the JBC Editorial Board. Edited by Karen Fleming

Crystallin proteins are a class of main structural proteins of the vertebrate eye lens, and their solubility and stability directly determine transparency and refractive power of the lens. Mutation in genes that encode these crystallin proteins is the most common cause for congenital cataracts. Despite extensive studies, the pathogenic and molecular mechanisms that effect congenital cataracts remain unclear. In this study, we identified a novel mutation in *CRYBB1* from a congenital cataract family, and demonstrated that this mutation led to an early termination of mRNA translation, resulting in a 49-residue C-terminally truncated CRYβB1 protein. We show this mutant is susceptible to proteolysis, which allowed us to determine a 1.2-Å resolution crystal structure of CRYβB1 without the entire C-terminal domain. In this crystal lattice, we observed that two N-terminal domain monomers form a dimer that structurally resembles the WT monomer, but with different surface characteristics. Biochemical analyses and cell-based data also suggested that this mutant is significantly more liable to aggregate and degrade compared to WT CRYβB1. Taken together, our results provide an insight into the mechanism regarding how a mutant crystalin contributes to the development of congenital cataract possibly through alteration of inter-protein interactions that result in protein aggregation.

Cataract, the opacity or light scattering of eye lens caused usually by the presence of abnormal crystallin protein aggregates, is the main cause of human blindness worldwide. It can be broadly divided into congenital and acquired cataract with congenital cataracts being one of the major causes of childhood blindness (1–3). Crystallin proteins are the major structural components of the vertebrate eye lens, and their solubility and stability are important for maintaining transparency and refractive power of the eye lens (1). They can be grouped into two families, the α-crystallins superfamily and β/γ-crystallins superfamily (4). Human β- and γ-crystallins share highly homologous amino acid sequences and common structural feature consisting of an N-terminal domain (NTD) and a C-terminal domain (CTD) connected by a short linker (5). These two domains are believed to evolve from a common single-domain protein ancestor by a series of gene duplication and fusion events (6, 7). Each domain consists of two similar highly stable Greek key motifs (GKMs) of about 40 amino acids, folding into a wedge-shaped β-sheet sandwich with a pseudo 2-fold symmetry (1). The main sequence difference between oligomeric β-crystallins and monomeric γ-crystallins is that only the former has long N-terminal extensions. The terminal extensions of β-crystallins are thought to be involved in the formation and stabilization of higher-order homo- or hetero-oligomer (8–10). The NTD and CTD of β/γ-crystallins arrange as an intermolecularly domain-swapped dimer or intramolecularly face-to-face dimer according to available structures in the Protein Data Bank (PDB) (1, 5, 6, 11–14). The structural integrity of the GKMs in β/γ-crystallins is vital for proper folding of protein, and mutations disrupting even one of the four GKMs would result in protein self-aggregation and precipitation, consistent with the phenotype of nuclear cataract (4, 15–17). Correspondingly, mutation of crystallin genes is a dominant pathogenic factor of the congenital cataract characterized by lens opacity (18).

β-crystallins can be further classified into basic and acidic groups. Basic crystallins (βB1, βB2, and βB3) contain both N- and C-terminal extensions, while acidic crystallins (βA1, βA2, βA3, and βA4) possess only N-terminal extension (5). CRYβB1 is a primary member of β-crystallins and comprises 9% of the total soluble crystallin proteins in the human lens (17, 18).

‡ These authors contributed equally to this work.
* For correspondence: Peng Gong, gongpeng@wh.iov.cn; Ya-Ping Tang, yptang12@gzhmu.edu.cn; Dao-Man Xiang, xiangdm35@126.com.

Clinically identified mutations (p.Q227X, p.Q223X, p.G220X, p.X253R, p.S228P, p.R233H, p.S93R, and p.S129R; each X denotes a translation termination codon and results in a C-terminal truncation of the protein) in CRYβB1 were found to be associated with the autosomal dominant cataract (4, 16, 17, 19–23). The C-terminal truncations lead to partial or complete deletion of the C-terminal extension or even affect the integrity of the CTD. For documented autosomal dominant congenital cataract-related CRYβB1 truncation-type mutations (p.Q227X, p.Q223X, and p.G220X), it has been proposed that the mutations disrupt the hydrophobic core of the CTD, leading the mutants to be insoluble and easier to aggregate, eventually causing cataract (4, 16, 20). However, the molecular mechanism underlying the pathogenesis of congenital cataract remains unclear. The three-dimensional structure and functional exploration of novel congenital cataract-related crystallin mutants could provide important references in elucidating the pathogenic mechanism and therefore beneficial for translational research efforts on curing congenital cataract.

In this study, we identified a previously unreported truncation-type mutation (nucleotide: c.612 C > A; amino acid: p.Y204X) in exon 6 of *CRYBB1*, a cataract-causative gene in a family with autosomal dominant congenital cataract. To explore the molecular mechanism of the pathogenic role, we solved a 1.20-Å resolution crystal structure of the mouse CRYβB1 truncated mutant Y202X (M_Y202X, the corresponding mutation of human p.Y204X in mouse, which also causes cataract in mouse). The structure only contains the CRYβB1 NTD, but forms WT CRYβB1 monomer-mimicking dimer in the crystal lattice. Furthermore, the mutant protein shows different structural and biochemical characteristics when compared to those in WT CRYβB1, with respect to oligomerization and proteolysis-related degradation. This work thus provides key information for further understanding of the molecular mechanism of this and similar pathogenic mutation.

Results

Identification of a novel CRYBB1 *gene mutation associated with cataract in a Chinese family*

In a three-generation Chinese family, four members suffered from nuclear cataract with an obvious hereditary tendency (Fig. 1, *A* and *B*). To identify the genetic lesions responsible for cataract in this family, whole exome sequencing (WES) was performed with genomic DNAs (gDNAs) from patients and their relatives. A heterozygous *CRYBB1* c.612 C > A mutation (NM_001887.4), which was not reported previously, was detected in all four patients (III:4, III:5, III:6, and III:7) and two unaffected members (II:3 and II:5) (Fig. 1*C*). This mutation transited the Y204 codon (TAC) to a termination codon (TAA), resulting in an earlier termination of translation, and thereby produced a truncation-type CRYβB1 Y204X mutant without C-terminal 49 amino acid residues (residues 204–252), in which the GKM IV and C-terminal extension was included (Fig. 2*A*). As II:3 and II:5 also carry the mutation but do not show cataract, we considered it a pathogenic mutation with incomplete penetrance. Next, to further confirm the pathogenicity of the mutation, we set out to explore the structure and biochemical properties of the truncated CRYβB1 protein.

Crystallization and structure determination of the mouse CRYβB1 Y202X mutants

The N-terminal hexahistidine-tagged human WT CRYβB1 (H_WT), human Y204X mutant (H_Y204X), mouse WT

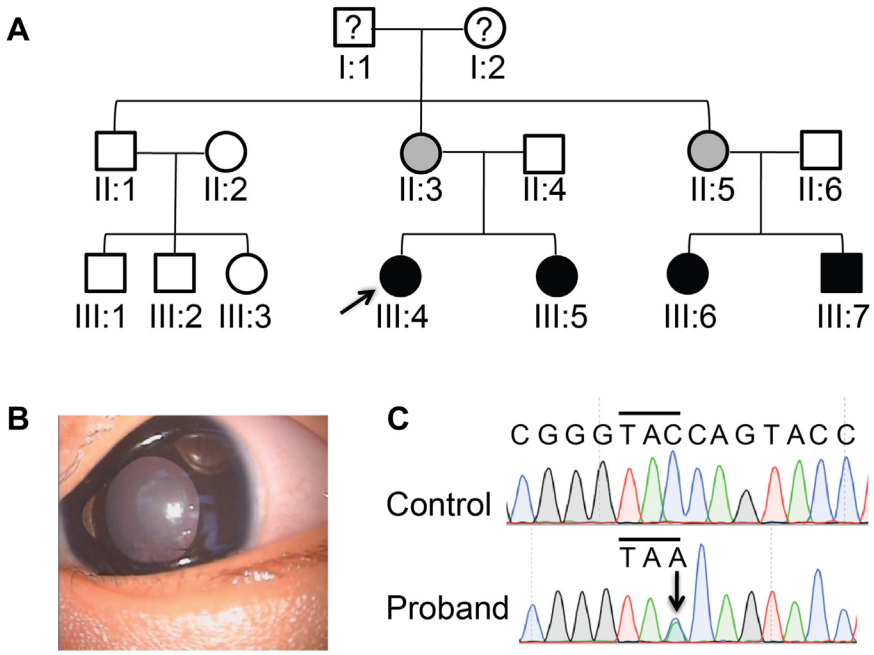

Figure 1. The mutation analysis of *CRYBB1* in a Chinese family with congenital nuclear cataract. *A*, pedigree of the family. *Squares* and *circles* indicate men and women, respectively. *Solid symbols* denote affected status and the *gray symbols* denote the normal people with the mutated gene. The proband is indicated with an *arrow*. The *question mark* denotes the person without sequencing. *B*, slit-lamp view of the lens of the proband shows the nuclear cataract. *C*, the DNA sequence chromatogram shows a C612A heterozygous mutation in *CRYBB1* indicated by an *arrow*.

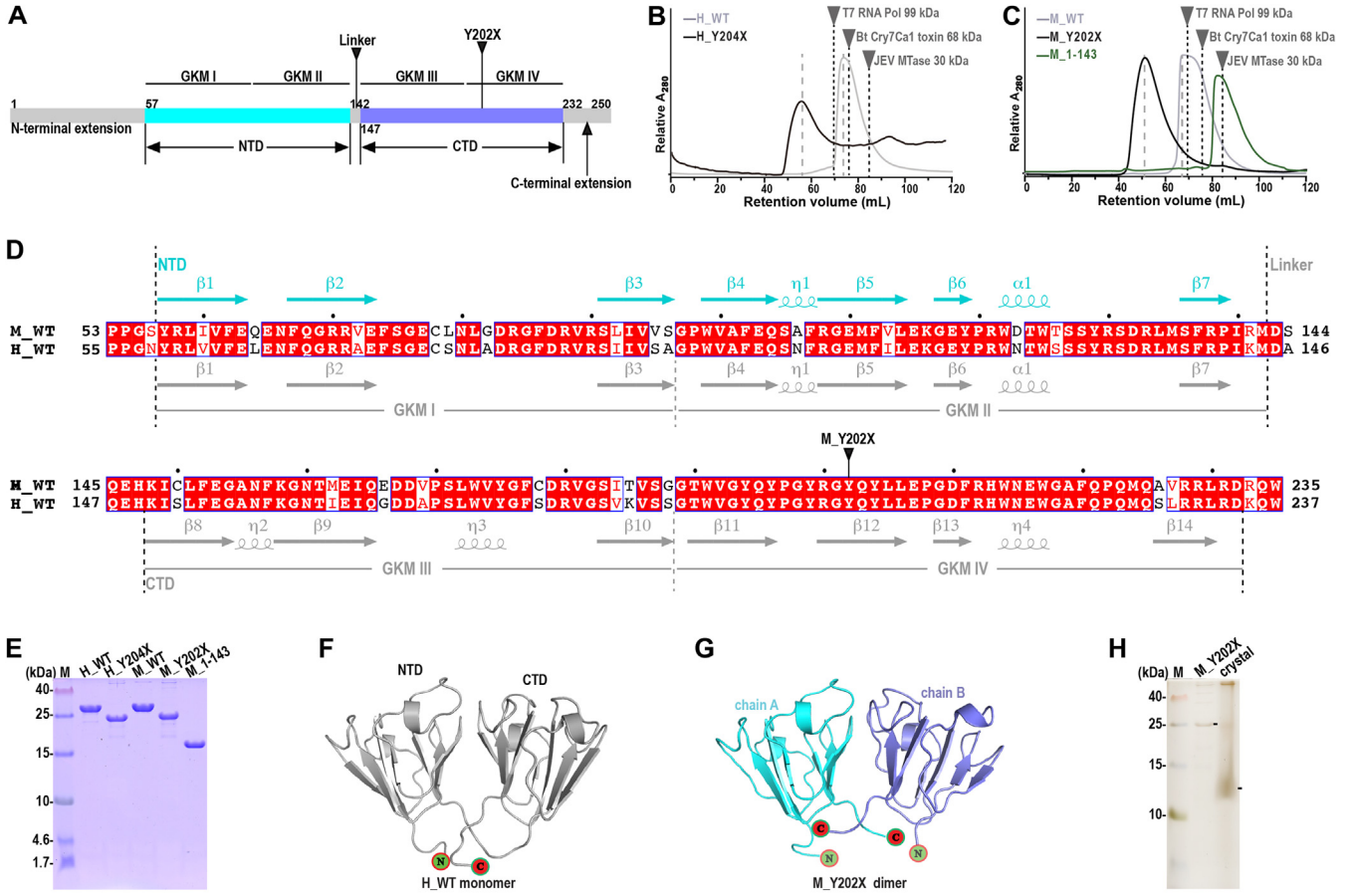

Figure 2. Overall structure of the CRYβB1 mutant. *A*, a *color-coded bar* defining structural elements of the mouse CRYβB1 WT protein. The N-terminal extension (*gray*), NTD (*cyan*), the linker (*gray*), the CTD (*slate*), and the C-terminal extension (*gray*) are individually color coded. The four Greek key motifs (GKMs) and the mutation site Y202X are indicated according to the residue region or site. *B* and *C*, samples derived from preceding ion-exchange chromatography were loaded onto the gel filtration (GF) column equilibrated with the GF buffer. H_Y204X and M_Y202X exhibited a HMW polymeric state in a superdex 200 GF column, while the H_WT and M_WT exist probably as a tetramer according to the retention volume. M_1-143 exists as a dimer. Empirical retention volumes and the molecular weights of three other proteins are indicated for comparison: JEV MTase, T7 RNA Pol, and Bt. *D*, the sequence alignment and secondary structure assignment of M_WT and H_WT. The boundaries of different domains and GKMs are marked by *vertical lines*. The mutation site M_Y202X is labeled. Helices and β-strands are indicated by *springs* and *arrows*, respectively. The identical amino acid residues of M_WT and H_WT were shown by *rectangles* filled with *red* backgrounds. NTD of M_WT is colored in *cyan* and all of H_WT in gray. *E*, the purified H_WT (29 kDa), H_Y204X (23 kDa), M_WT (29 kDa), M_Y202X (23 kDa), and M_1-143 (16.5 kDa) proteins were detected in Tris-Tricine-SDS-PAGE. *F*, the monomer structure of the human CRYβB1 H_WT protein (PDB: 1OKI). The whole structure is colored in *gray*70. *G*, the M_Y202X dimer structure. Chain A is colored in *cyan*, and chain B in *slate*. The N and C termini are shown by *circles* filled with *green* and *red* backgrounds, respectively. *H*, the M_Y202X crystal analysis in Tris-Tricine-SDS-PAGE. The M_WT, M_Y202X, and crystal bands are indicated by *short lines*. *Bacillus thuringiensis*; CTD, C-terminal domain; GKMs, Greek key motifs; HMW, high molecular weight; JEV MTase, Japanese encephalitis virus methyltransferase; NTD, N-terminal domain; PDB, Protein Data Bank; T7 RNA Pol, T7 RNA polymerase.

CRYβB1 (M_WT), mouse CRYβB1 mutant (M_Y202X), and mouse 1 to 134 (M_1–143) were purified and then subjected to crystallization trials (Fig. 2, *A–E*). Block shape crystals of the M_Y202X were obtained in initial crystallization screening after one week while crystallization of the other four proteins was unsuccessful. The best M_Y202X single crystals were at least 0.1 mm in each dimension and allowed us to obtain an X-ray diffraction data set at 1.20-Å resolution. As the sequence identity between the mouse and human CRYβB1 proteins is 80% with 95% residue coverage (Fig. 2*D*), we chose a human CRYβB1 structure (1) as the template to generate molecular replacement (MR) search models (Fig. 2*F*). Successful MR was achieved with a modified NTD model that contains 88 residues (human CRYβB1 residues 54–141). Silver staining of the crystal samples indicated that the molecular weight of the crystallized M_Y202X protein (23 kDa) is in the range of 10 to 15 kDa, suggesting degradation of the protein (Fig. 2*H*). R_{work}/R_{free} values after the initial round of structure refinement were 0.204/0.216 (Table 1). After reiterative model building and refinement trials, the final structure model had R_{work}/R_{free} values of 0.185/0.194 with 91 residues resolved (Table 1). The crystallographic asymmetric unit comprises a dimer with nearly identical conformation. The RMSD value for all α-carbon atoms is 0.6 Å with 96% coverage (chain A as the reference), and the arrangement of the two NTD monomers mimics one WT protein monomer. The RMSD value for all superimposable α-carbon atoms is 1.1 Å with 93% coverage (mutant structure as the reference, Fig. 2*G*). Despite the observed dimeric form in the crystal lattice, the M_Y202X primarily exists as high molecular weight (HMW) species in gel filtration (GF) chromatography (Fig. 2*C*).

Table 1
X-ray diffraction data processing and structure refinement parameters

Parameter	Value
Data processing[a]	
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	39.9, 48.4, 113.3
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	50.00–1.20 (1.24–1.20)[b]
R_{merge}	0.058 (0.482)
R_{meas}	0.061 (0.543)
$CC_{1/2}$	0.999 (0.846)
I/σI	35.8 (3.1)
Completeness (%)	98.1 (84.5)
Redundancy	10.4 (4.6)
Refinement	
Resolution (Å)	1.20
No. of unique reflections	68,183
R_{work}/R_{free}	0.185/0.194
No. atoms	
Protein	1507
Ligand/ion/water	−/−/187
B-factors (Å^2)	
Protein	18.1
Ligand/ion/water	−/−/30.0
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.890
Ramachandran statistics[c]	89.3/10.7/0.0/0.0

[a] One crystal was used for this structure.
[b] Values in parentheses are for the highest-resolution shell.
[c] Values are in percentage and are for the most favored, additionally allowed, generously allowed, and disallowed regions in Ramachandran plots, respectively.

Overall structure of the mouse CRYβB1 M_Y202X mutant and a comparison with WT protein structure

The two M_Y202X NTD monomers pair in a symmetrical manner very similar to that observed in the CRYβB1 H_WT crystallin, where its NTD and CTD (sharing ∼33% sequence identity in mouse) are related by a pseudo 2-fold symmetry (Fig. 2, *A*, *D*, *F*, and *G*). The M_Y202X NTD structures are almost identical to the NTD of the WT protein (RMSD values for 87–88 superimposable α-carbon atoms: 0.5–0.6 Å, chain A of the mutant structure as the reference) (Fig. 2, *F* and *G*). The intermolecular interaction interface between the M_Y202X NTD monomers is highly similar to the intramolecular NTD-CTD interaction interface of the H_WT. Both interfaces are featured by hydrogen bonding, charged, and hydrophobic interactions (Fig. 3). Two hallmark pairs of arginine and glutamic acid form salt bridges in M_Y202X NTD dimer and the H_WT monomer, respectively, while a notable difference is that a pair of R130 residues (one from each NTD monomer) interacts with each other at the edge of the interface in the mutant structure (Fig. 3).

Surface characteristics of the mouse CRYβB1 mutant Y202X

To compare the surface characteristics between the M_Y202X NTD dimer (M1 and M2) and the WT monomer

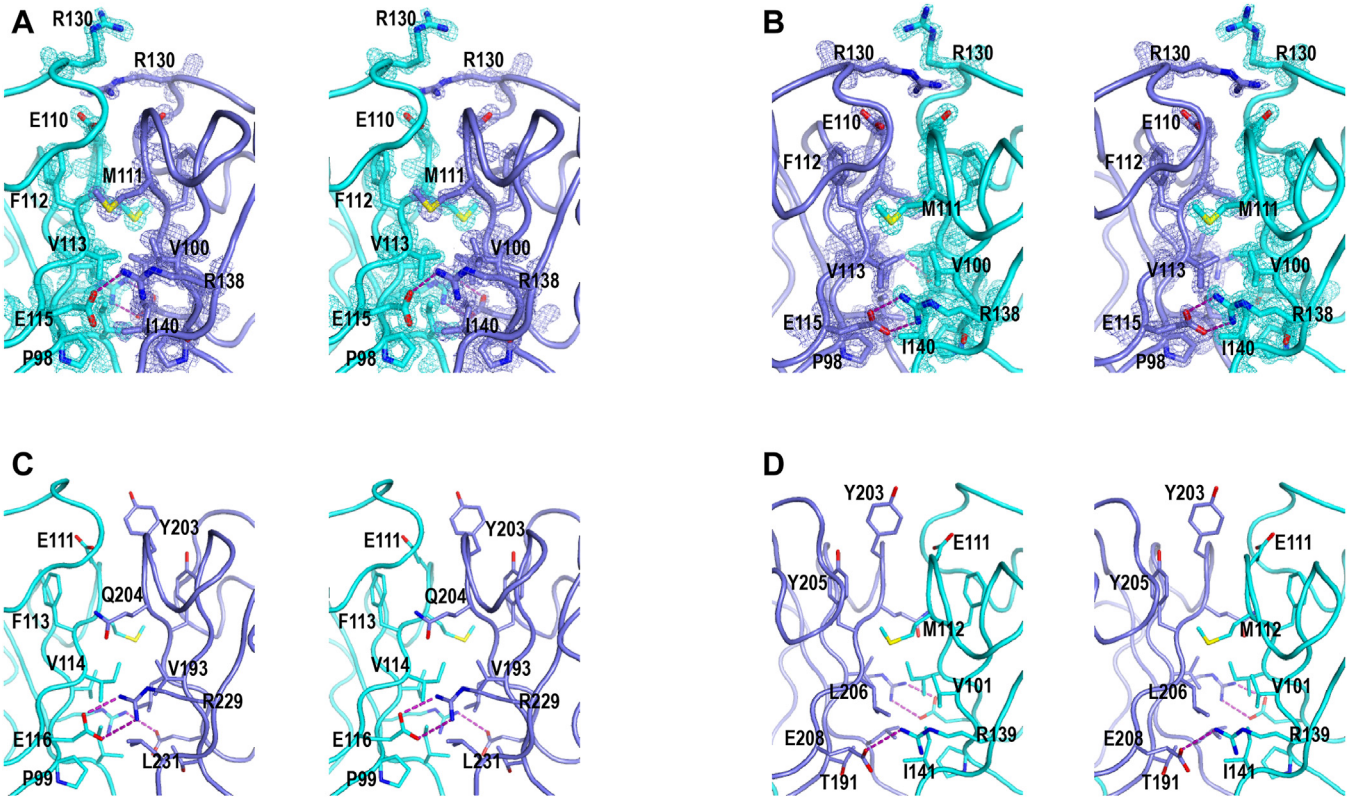

Figure 3. The interface comparison of the CRYβB1 M_Y202X dimer and M_WT monomer. *A* and *B*, the intermolecular interface of the CRYβB1 M_Y202X dimer. The viewing angles of the two panels are correlated by 180° through the *y*-axis. *C* and *D*, the intramolecular interface of the CRYβB1 M_WT monomer. The viewing angles of the two panels are correlated by 180° through the *y*-axis. Stereo-pair images of the composite SA omit electron density map (contoured at 1.5 σ, see Experimental procedures) overlaid onto structural models around the intermolecular interface of M_Y202X dimer. The M_Y202X dimer is colored as Figure 2*G*. The NTD and CTD of WT protein shown in thinner representations are also colored in *cyan* and *slate*, respectively. The residues located in the exterior of the stereo-pair structures are shown as *sticks* and labeled. The hydrogen bonds formed by arginine and glutamic acids are shown as *purple dashed lines* (distance range: 2.5–3.5 Å). CTD, C-terminal domain; NTD, N-terminal domain; SA, simulated-annealing.

(NTD and CTD), a M_WT protein model was obtained using the PhyRe2 server (24). Though the M_Y202X NTD dimer can well mimic the WT protein monomer in shape, it shows different electrostatic features in the solvent exposed surface of the second NTD monomer (M2, corresponding to the WT CTD in the comparison), with M2 displaying overall opposite electrostatic potentials in the top, side, and bottom regions (Fig. 4). These differences could help understand the low solubility and aggregation of this mutant.

The mutants are more liable to aggregate than WT proteins

In order to further explore the effects brought by the truncation-type mutation regarding solution properties, we analyzed M_WT, M_Y202X, H_WT, and H_Y204X by several biochemical methods. The GF chromatography data showed that the both mutants are HMW polymers, while both WT protein are low molecular weight oligomers according to their retention volumes in a superdex 200 GF column (Fig. 2, *B* and *C*). The dynamic light scattering data indicated that the hydrodynamic sizes of both mutants in PBS are larger than those of the WT proteins at 25 °C (Fig. 5*A*). Moreover, the negative stain EM data also showed that the sizes of the mutant particles are inhomogeneous and much larger than the WT proteins (Fig. 5*B*). To test whether the mutant protein forms aggregates in cells likewise, we constructed two vectors, N-myc-H_WT and N-myc-H_Y204X, to express H_WT and H_Y204X in human lens epithelial (HLE) cells HLE-B3, respectively. Through immunofluorescence analysis, we found that the H_Y204X mutant indeed formed abnormal large aggregates in cells (Fig. 5*C*). The propensity of the mutants to form HMW aggregates is compatible with the clinical observation that truncation-type mutations lead to the human congenital cataract.

The mutants are more sensitive to the trypsin than WT proteins

The M_Y202X and H_Y204X mutants are susceptible to degradation in the process of protein purification. In order to assess the degradation behaviors of the mutants in solution, trypsin proteolysis analyses were performed. Both mutants were degraded to higher degrees than the WT proteins under the same proteolysis condition, indicated by higher amount of intermediate products in mutants (Fig. 6, lanes 4–15 for M_WT, lanes 19–30 for M_Y202X, lanes 34–45 for H_WT, lanes 49–60 for H_Y204X). The main degradation intermediate (indicated by solid triangles in Fig. 6) was identified to be CRYβB1 proteins by liquid chromatography-tandem mass spectrometry (LC-MS/MS) and they all included the NTD region (Table S1). These data suggest that both mutations lead to enhanced protein sensitivity to trypsin proteolysis in solution.

Discussion

According to the public Human Gene Mutation Database and previous reports, 17 *CRYBB1* mutations have been associated with the congenital cataract (4, 16, 17, 19–23, 25, 26). Of them, eight mutations (including three truncation-type mutations) are within CRYβB1 C-terminal region (CTD and C-terminal extension). Previous functional investigation of these CRYβB1 C-terminal region mutations mainly focused on protein solubility, thermal stability, and oligomerization properties, but is not sufficient to elucidate mechanistic details. In this study, we clinically identified a novel cataract-causing CRYβB1 mutation p.Y204X, allowing us to study the mechanism of truncation-type mutations from a different mutation site. Besides dynamic light scattering, EM and cell-based data that support alteration in oligomerization propensity caused by the mutation, the M_Y202X NTD crystal structure and trypsin proteolysis analysis together suggest that this mutation disrupts structural integrity of CRYβB1 CTD and could result in complete degradation of the C-terminal region. As suggested by our structural analysis, new NTD-dictating interactions can be established between degraded CRYβB1. We therefore propose a working model for understanding the mechanism of action for CRYβB1 truncation-type mutation (Fig. 7). In contrast to WT protein that forms dimer through

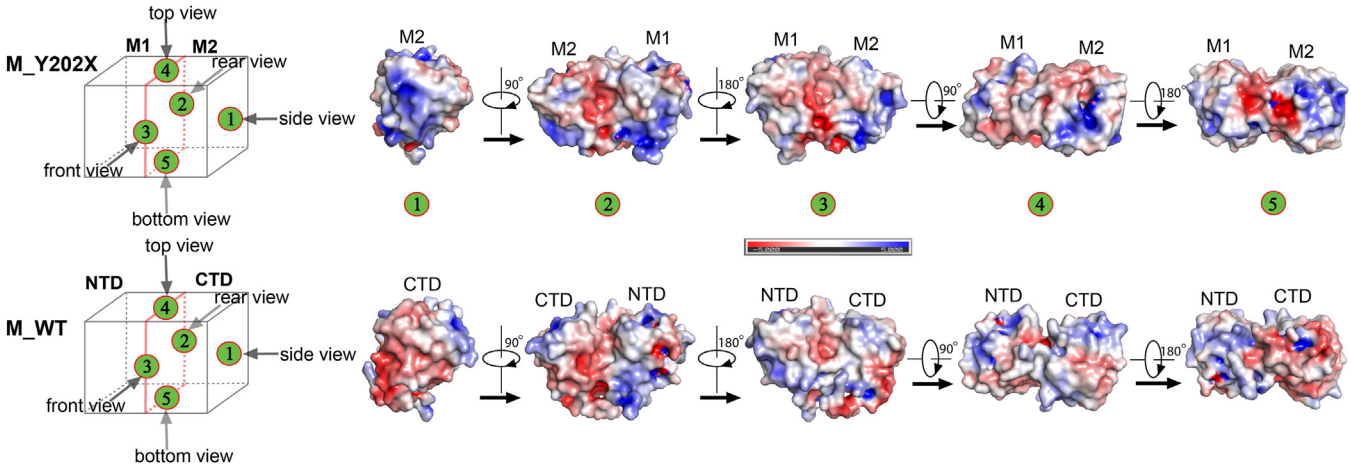

Figure 4. A surface characteristic comparison of the CRYβB1 mutant M_Y202X and M_WT protein. A diagram showing different views of M_Y202X or M_WT is drawn to distinguish the different surfaces more easily. The two domains (M1 and M2, NTD and CTD) are divided by *red solid* and *dotted lines*. The 1, 2, 3, 4, and 5 in *circles* filled with *green* and *red* backgrounds denote the side, rear, front, top, and bottom views, respectively. The *top panel* is the M_Y202X and the *bottom panel* is the M_WT. CTD, C-terminal domain; NTD, N-terminal domain.

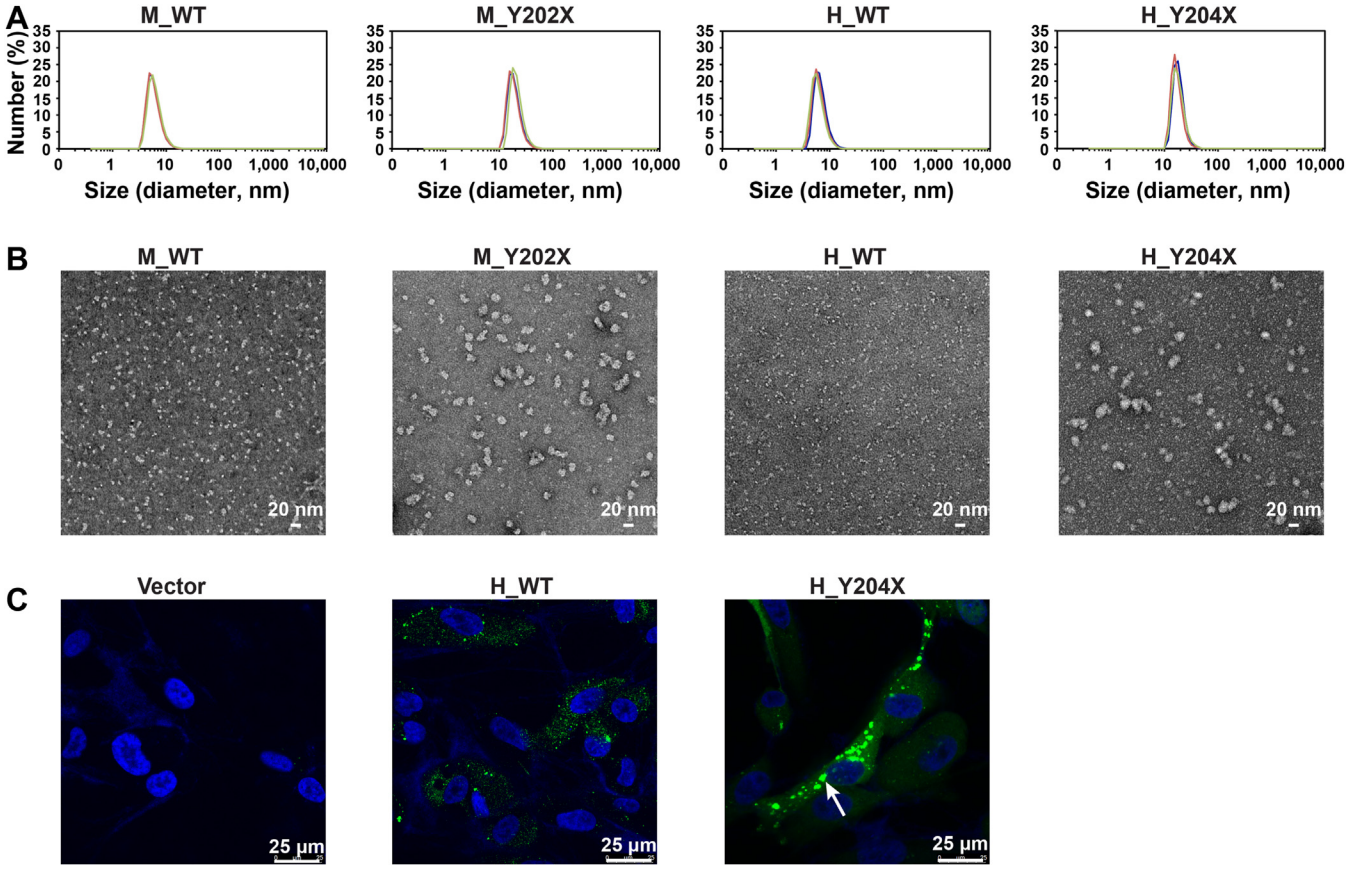

Figure 5. The aggregation formation analysis of CRYβB1 WT and mutated proteins. *A*, distribution of protein hydrodynamic diameters in PBS measured by DLS. The peaks of the M_Y202X and H_Y204X are larger than the WT proteins. *B*, representative negative staining EM micrographs of the CRYβB1 proteins. The scale bar represents 20 nm. *C*, the H_Y204X mutant protein forms abnormal large aggregates in HLE-B3 cells. The myc-H_WT or myc-H_Y204X protein was indicated by anti-myc (*green*). The abnormal large aggregate was labeled by *white arrow*. Nucleus was indicated by DAPI (*blue*). The scale bar represents 25 μm. DAPI, 4′,6-diamidino-2-phenylindole; DLS, dynamic light scattering.

intermolecular interactions involving intact NTD and CTD, truncation-type mutants disrupt the structural integrity of CTD either directly or through degradation, thus altering the intermolecular interaction profiles and may lead to formation of HMW aggregates (Fig. 7). As indicated by the M_Y202X NTD structure, complete degradation of the C-terminal region can promote NTD dimer formation. One NTD dimer can well mimic one WT monomer in shape but not in surface properties such as electrostatic features.

Two primary models have been proposed for mutation-derived cataract formation (27–29). The first model is featured by mild mutations that do not significantly change protein properties. However, long-term impact from adverse environmental factors may eventually cause protein denaturation and aggregation that affects light scattering and leads to cataract. In the second model, protein structure is disrupted to a great extent by severe mutations that cause congenital cataracts. The truncation-type mutations, including CRYβB1 p.Y204X mutation in this study, belong to the severe mutation category. Here, we investigated the molecular properties of congenital cataract-related CRYβB1 p.Y204X mutation in comparison with the WT protein and their counterparts in mouse. Our data, in particular those obtained through crystallography, trypsin proteolytic analysis, and cell-based data, suggest that this mutation may change the intermolecular interactions between CRYβB1 protein and other crystallins, either through the intact mutant or the degraded NTD dimer that mimics WT monomer (Fig. 7). Such mutation-derived alteration of intermolecular interactions could lead to malfunctioning of lens cells or disruption of crystallin lens microstructures. Our work thus provides an important reference in understanding the structure–function relationship of crystallins and the connection between crystallin mutation and cataract formation.

Experimental procedures

Ethics

This study was conducted in accordance with the Declaration of Helsinki. The study protocols were approved by the Human Ethics Committee of the Guangzhou Women and Children's Medical Center. Written informed consent was obtained from each participant or their legal custodians.

Whole exome sequencing

For WES, gDNA was extracted from the peripheral blood of all subjects and sonicated into DNA fragments of 150 to 300 bp. WES was performed at the Beijing Genomics Institute as

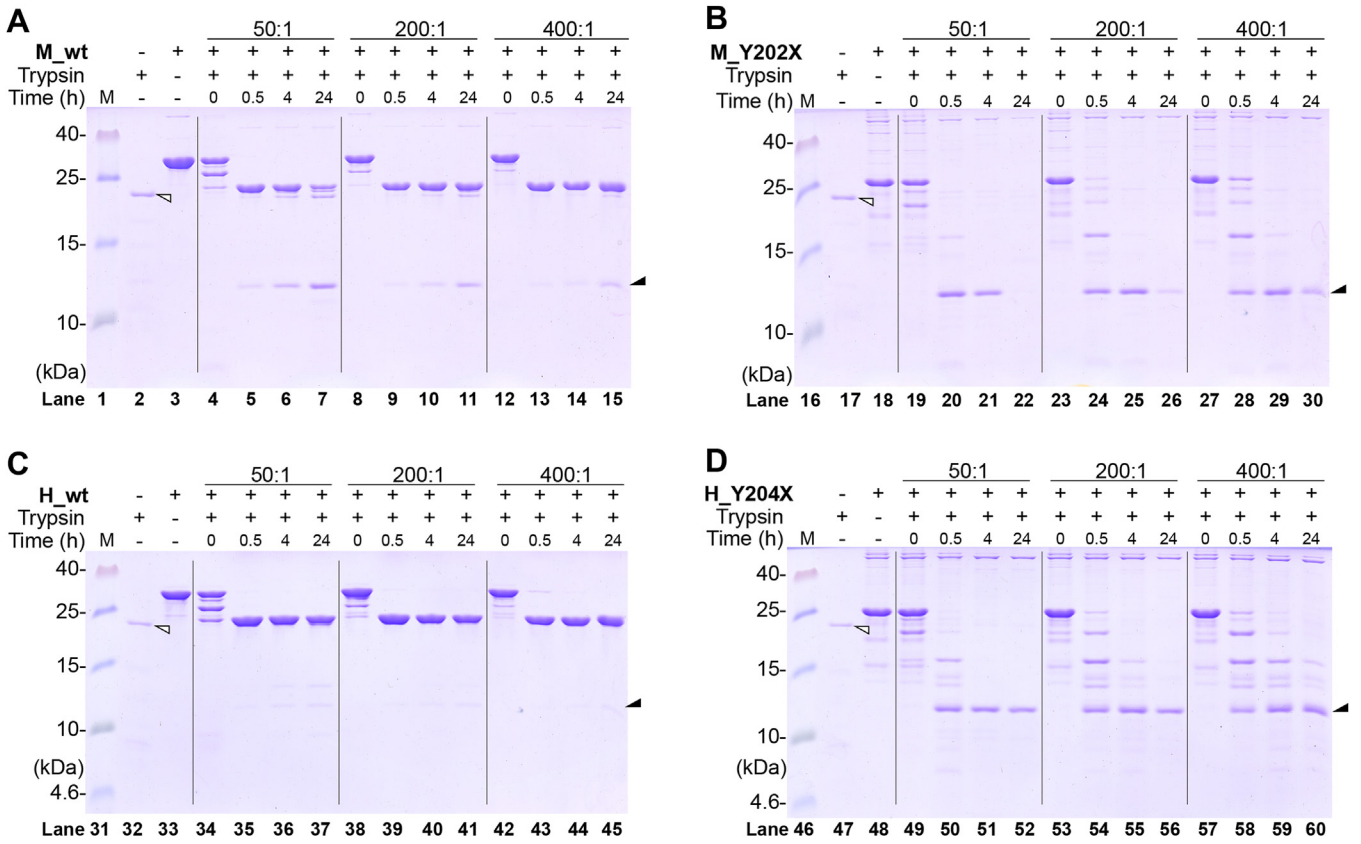

Figure 6. The trypsin proteolysis analyses of the CRYβB1 proteins. *A–D*, The four CRYβB1 proteins (M_WT (*A*) in lane 1–15, M_Y202X (*B*) in lane 16–30, H_WT (*C*) in lane 31–45, H_Y204X (*D*) in lane 46–60) were treated with protein and trypsin at different ratio (protein (wt.) to trypsin (wt.), 50:1, 200:1, and 400:1, divided by *black thin line*) in a series of time points. The samples at time point 0 h were performed through adding the loading buffer immediately after the trypsin mixing with the CRYβB1 proteins. The trypsin bands were labeled by *empty triangles* and the dominant degradation intermediate bands of CRYβB1 proteins were labeled by *solid triangles* and identified by LC-MS/MS.

previously described (30). The raw data were collected using Illumina Base Calling software (https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html) (bcl2fastq). The human genome assembly hg19 (GRCh37) was used as the reference sequence. The Genome Analysis Toolkit (v3.3.0) (https://software.broadinstitute.org/gatk/) and ANNOVAR software (https://annovar.openbioinformatics.org/en/latest/) were employed to detect and annotate the variants, respectively. Then, all variants were filtered following a pipeline: (1) exclusion of variants with a frequency greater than 1% in any of the four databases (1000g_all, esp6500siv2_all, gnomAd_ALL, and gnomAD_EAS); (2) exclusion of variants that were not in the coding (exonic) region or splicing region (splicing site ± 10 bp); (3) exclusion of synonymous SNPs that were not predicted by dbscSNV to affect splicing; and (4) retention of variants that were predicted by at least two of four prediction tools (SIFT, PolyPhen, MutationTaster, and CADD) to be deleterious and variants that are predicted to affect splicing.

Variant verification

Sanger sequencing was performed to validate the pathogenic variant identified by WES. gDNA was used as a template. The following primers were used: 5′-TACCATGCACAGG CAACATGC-3′ (forwards) and 5′-TAGCAGAGTGAGG TGTGGACTC-3′ (reverse). The qualified PCR products were sent to Shanghai Sangon Biotech for sequencing. *CRYBB1* reference sequence: NM_001887.4.

Plasmid construction, protein expression, and purification

The human CRYβB1 H_WT, human CRYβB1 H_Y204X, mouse CRYβB1 M_WT, mouse mutant M_Y202X (codon optimized), and M_1-143 (mouse 1–143) genes were cloned into the pET26b vector. The restriction enzyme sites used for M_WT were *Bam* HI/*Xho* I, and M_Y202X, M_1-143, H_WT and H_Y204X were achieved through a site-directed ligase-independent mutagenesis method (31). The resulting plasmids were transformed into *Escherichia coli* strain BL21(DE3) for the expression of the proteins. The cells were cultured in LB medium containing 50 μg/ml for kanamycin at 37 °C until the absorbance at 600 nm (A_{600}) reached 0.4 to 0.6. Then the cultures were induced with IPTG at a concentration of 0.5 mM at 16 °C overnight. The cells were harvested by centrifugation at 6740*g* for 15 min in an F10S × 1000 rotor (Thermo Fisher Scientific) and stored at −80 °C. The CRYβB1 proteins purification procedures were modified from previously reported methods with glycerol removed from the GF buffer (300 mM NaCl, 5 mM Tris pH 7.5) in the third chromatography step (32). The lysates were in turn loaded onto nickel-charged HisTrap HP column (GE HealthCare), HiTrap SP/Q HP column (GE HealthCare) and Superdex 200 gel filtration column

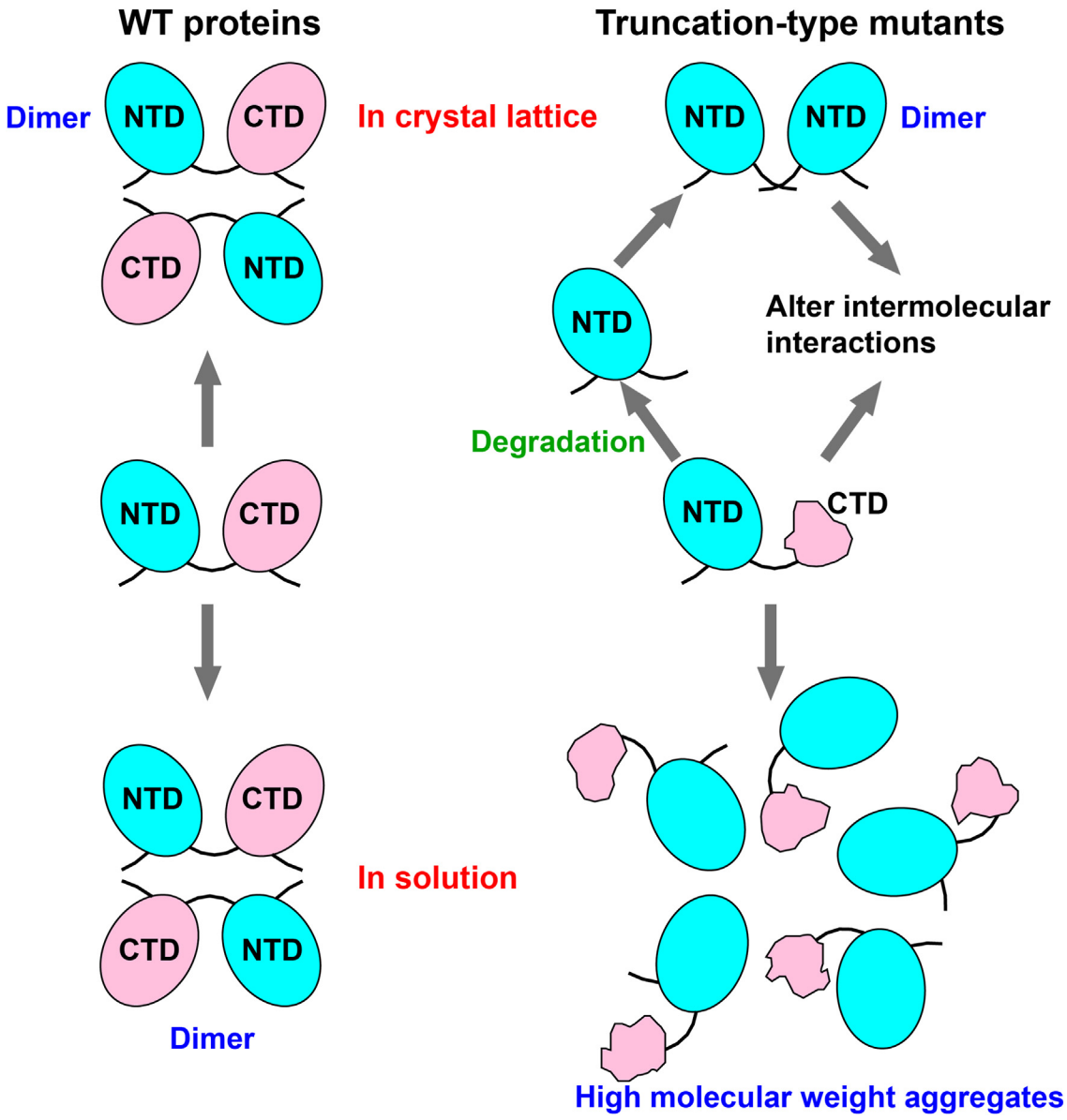

Figure 7. Proposed mechanism of action for *CRYBB1* truncation-type mutations. *Left*: WT CRYβB1 WT proteins form dimers both in crystal lattice and in solution. *Right*: CRYβB1 truncation-type mutants (*e.g.*, H_Y204X) form HMW aggregates in solution. These mutants may be susceptible to degradation and become NTD-only variants as captured in CRYβB1 H_Y204X crystal lattice. The NTD only variant of CRYβB1 can form dimer, and together with the original form of the mutant, may alter intermolecular interactions of native crystallin proteins. The NTD and CTD of CRYβB1 are colored in *cyan* and *pink*, respectively, and the N- and C-terminal extensions are shown as *black thin lines*. WT proteins: M_WT or H_WT; truncation-type mutants: M_Y202X or H_Y204X. CTD, C-terminal domain; HMW, high molecular weight; NTD, N-terminal domain.

(GE HealthCare). M_WT and M_Y202X were loaded onto HiTrap Q HP column (25 mM Tris pH 8.5, 0.1 mM EDTA, 20% (v/v) glycerol, and 0/1 M NaCl) and M_1-143, H_WT and H_Y204X were loaded onto HiTrap SP HP column (25 mM Mes pH 6.0, 0.1 mM EDTA, 20% (v/v) glycerol, and 0/1 M NaCl). The NaCl concentrations of H_WT and M_WT loading onto HiTrap SP/Q HP column were 20 mM and 50 mM for H_Y204X, M_Y202X, and M_1-143. The purified protein was supplemented with 5 mM tris-(2-carboxyethyl) phosphine (pH 7.5, adjusted by KOH) and concentrated to a final concentration of about 3 to 20 mg/ml, flash frozen with liquid nitrogen, and stored at −80 °C as 10 to 20 μl aliquots for single use. The extinction coefficient of CRYβB1 proteins were calculated based on the amino acid sequences using the ExPASy ProtParam program (http://www.expasy.ch/tools/protparam.html).

Crystallization, data collection, and structure determination

The crystallization screening of CRYβB1 proteins was performed at 289 K by sitting drop vapor diffusion. The M_Y202X plate-like crystal yielding the final X-ray diffraction dataset was obtained by mixing 0.3 μl of 10 mg/ml M_Y202X, 0.3 μl of a reservoir solution of 0.1 M sodium malonate (pH 6.0), and 12% (w/v) PEG 3350. Crystals were soaked in the precipitant solution supplemented with 26.7% (vol./vol.) glycerol as a cryoprotectant before being flash-cooled and stored in liquid

nitrogen. The X-ray diffraction data were collected at the Shanghai Synchrotron Radiation Facility beamline BL17U1 (wavelength = 0.9792 Å, temperature = 100 K) (33). Reflections were integrated, merged, and scaled using HKL2000 (Table 1) (34). The initial structure was obtained using the MR program PHASER (35) with a truncated human CRYβB1 crystal structure (PDB entry 1OKI) (1) as a search model with the CTD removed. Manual model rebuilding and refinement was completed by Coot (https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/) (36) and PHENIX (https://phenix-online.org), respectively (37). Noncrystallographic symmetry was applied to both chains in the asymmetric unit in initial rounds of refinement and was released in later rounds. The Ramachandran statistics are 89.3%, 10.7%, 0.0%, and 0.0% for favored, allowed, generously allowed, and disfavored regions, respectively. The 3500 K composite simulated-annealing omit $2F_o–F_c$ electron density maps were generated by Crystallography and NMR Systems (CNS) (38). All structure superimpositions were performed using the maximum likelihood structure superpositioning program THESEUS (39). The PDB entry used for superpositioning analysis with CRYβB1 Y202X is 1OKI (CRYβB1 WT) (1).

Measurement of hydrodynamic size

In order to compare the aggregation tendency of the CRYβB1 proteins in solution, the H_WT, H_Y204X, M_WT, and M_Y202X protein samples in PBS were centrifuged at 15,871*g* in a 5424/5424R rotor (Eppendorf) for 20 min at 25 °C. The supernatants of the samples were used for the measurement of the hydrodynamic size of the protein particles. Measurements were performed on a Zetasizer Nano ZS90 (Malvern Panalytical Ltd) and the parameters were automatically optimized. Each sample was measured three times.

Negative staining analysis

For negative staining assays, the CRYβB1 proteins were diluted in the GF buffer to 0.02 mg/ml. Ten microliter of protein sample was loaded onto a glow-discharged carbon-coated copper grid and stained with 3% uranyl acetate. The prepared grids were examined using an FEI Tecnai 20 TEM (operated at 200 kV) equipped with a Gatan UltraScan 894 charge-coupled device camera (CCD). The transmission electron microscopy images were trimmed using Photoshop software (https://www.adobe.com/cn/products/photoshop.html) without any brightness/contrast adjustment.

Construction of expression plasmids and transfection of HLE-B3 cells

The coding sequence corresponding to the full-length WT *CRYBB1* (NM_001887) ORF was amplified from a commercial human *CRYBB1* ORF clone (F121325, Youbao Biological). The complementary DNA was inserted between the *Sal* I and *Bam* HI restriction sites on the pcDNA3.1-N-myc vector, and then sequenced. The generated construct, hereafter N-myc-H_WT, expressed an N-terminal-myc tagged WT human CRYβB1 under cytomegalovirus (CMV) promoter. To generate the N-myc-H_Y204X construct, N-myc-H_WT was mutated with an M5 Site-Directed Mutagenesis Kit (Mei5bio). The following primers were used: forward 5′-GCGGGTAACAGTACCT CCTAGAGC-3′ and reverse 5′-ACTGTTACCCGCGGTGA-TAGC-3′.

HLE-B3 is a human lens epithelial cell line. The empty vector, the N-myc-H_WT construct and the N-myc-H_Y204X construct were transfected into HLE-B3 cells with Lipofectamine 2000 (Invitrogen). Under 5% CO_2 at 37 °C, HLE-B3 cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% (v/v) fetal bovine serum (Gibco), 100 U/ml penicillin G, and 100 μg/ml streptomycin.

Immunofluorescence and confocal imaging

Transfected cells were fixed with 4% paraformaldehyde for 25 min and washed three times in PBS. Then, the cells were permeabilized in 0.5% Triton X-100/PBS for 10 min and subsequently blocked in 5% normal goat serum in PBS for 1 h at room temperature. The cells were then incubated with mouse anti-myc (1:1000, Cell Signaling Technology) at 4 °C overnight. The secondary antibodies used were Alexa Fluor 488 goat anti-rabbit IgG (1:1000). Finally, the cells were counterstained with DAPI (Vector Lab). Confocal images were captured on a Leica TCS SP8 laser scanning confocal microscope.

Trypsin proteolysis assays

A 20 μl reaction mixture containing protein samples (0.5 mg/ml) and trypsin at 50:1 (0.01 mg/ml), 200:1 (0.0025 mg/ml) and 400:1 (0.00125 mg/ml) in 25 mM Tris-HCl (pH 8.0) was incubated at 25 °C for 0, 0.5, 4, and 24 h before being boiled with an equal volume of Tris-Tricine-SDS-PAGE loading buffer (40 mM Tris-HCl, pH 8.0, 40% (v/v) glycerol, 20 mM DTT, 4% (w/v) SDS, 0.02% (w/v) bromophenol blue). The samples were analyzed by Tris-Tricine-SDS-PAGE. The dominant degradation intermediate bands at 12 to 13 kDa were identified by LC-MS/MS at Shanghai Sangon Biotech.

Data availability

All data acquired are available upon request.

Accession number

The atomic coordinates and structure factor files have been deposited in the Protein Data Bank under accession code 8H0R.

Supporting information—This article contains supporting information.

Acknowledgments—We thank synchrotrons SSRF (beamline BL17U1/BL10U2 Shanghai, China) for access to beamline, Dr Jiqin Wu, Xiang Fang, Qiaojie Liu, Tianying Nong, and Caixia Xian for laboratory assistance.

Author contributions—X. J., M. Z., X. L., P. W., L. S., and B.-Y. Z. investigation; X. J., M. Z., B.-Y. Z., Y.-P. T., D.-M. X., and P. G. visualization; X. J., M. Z., B.-Y. Z., Y.-P. T., D.-M. X., and P. G. writing–original draft; X. J., M. Z., P. W., L. S., B.-Y. Z., Y. X., Y.-P. T., D.-M. X., and P. G. writing–review and editing; X. J., M. Z., Y.-P. T., and P. G. funding acquisition; M. Z., X. L., P. W., and D.-M. X. resources; M. Z. project administration; Y. X., Y.-P. T., D.-M. X., and P. G. conceptualization.

Funding and additional information—This study was supported by the National Natural Science Foundation of China (82101102 to X. J.), the Key Biosafety Science and Technology Program of Hubei Jiangxia Laboratory (JXBS001 to P. G.), the National Key Research and Development Program of China (2021YFC2701800 and 2021YFC2701801 to Y. X.); and the Natural Science Foundation of Guangdong Province, China (2022A1515012506 to M. Z.).

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: CTD, C-terminal domain; gDNA, genomic DNA; GF, gel filtration; GKM, Greek key motif; HMW, high molecular weight; MR, molecular replacement; NTD, N-terminal domain; PDB, Protein Data Bank; WES, whole exome sequencing.

References

1. Van Montfort, R. L., Bateman, O. A., Lubsen, N. H., and Slingsby, C. (2003) Crystal structure of truncated human betaB1-crystallin. *Protein Sci.* **12**, 2606–2612
2. Zhao, W. J., and Yan, Y. B. (2018) Increasing susceptibility to oxidative stress by cataract-causing crystallin mutations. *Int. J. Biol. Macromol.* **108**, 665–673
3. Churchill, A., and Graw, J. (2011) Clinical and experimental advances in congenital and paediatric cataracts. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 1234–1249
4. Rao, Y., Dong, S., Li, Z., Yang, G., Peng, C., Yan, M., *et al.* (2017) A novel truncation mutation in CRYBB1 associated with autosomal dominant congenital cataract with nystagmus. *Mol. Vis.* **23**, 624–637
5. Xi, Z., Whitley, M. J., and Gronenborn, A. M. (2017) Human betaB2-Crystallin forms a face-en-face dimer in solution: an integrated NMR and SAXS study. *Structure* **25**, 496–505
6. Norledge, B. V., Mayr, E. M., Glockshuber, R., Bateman, O. A., Slingsby, C., Jaenicke, R., *et al.* (1996) The X-ray structures of two mutant crystallin domains shed light on the evolution of multi-domain proteins. *Nat. Struct. Biol.* **3**, 267–274
7. Clout, N. J., Basak, A., Wieligmann, K., Bateman, O. A., Jaenicke, R., and Slingsby, C. (2000) The N-terminal domain of betaB2-crystallin resembles the putative ancestral homodimer. *J. Mol. Biol.* **304**, 253–257
8. Ajaz, M. S., Ma, Z., Smith, D. L., and Smith, J. B. (1997) Size of human lens beta-crystallin aggregates are distinguished by N-terminal truncation of betaB1. *J. Biol. Chem.* **272**, 11250–11255
9. Werten, P. J., Lindner, R. A., Carver, J. A., and de Jong, W. W. (1999) Formation of betaA3/betaB2-crystallin mixed complexes: involvement of N- and C-terminal extensions. *Biochim. Biophys. Acta* **1432**, 286–292
10. Srivastava, K., Gupta, R., Chaves, J. M., and Srivastava, O. P. (2009) Truncated human beta B1-crystallin shows altered structural properties and interaction with human beta A3-crystallin. *Biochemistry* **48**, 7179–7189
11. Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P. F., Mahadevan, D., *et al.* (1990) X-Ray-analysis of beta-B2-crystallin and evolution of oligomeric lens proteins. *Nature* **347**, 776–780
12. Kumaraswamy, V. S., Lindley, P. F., Slingsby, C., and Glover, I. D. (1996) An eye lens protein-water structure: 1.2 angstrom resolution structure of gamma B-crystallin at 150K. *Acta Crystallogr. D* **52**, 611–622
13. Wright, G., Basak, A. K., Wieligmann, K., Mayr, E. M., and Slingsby, C. (1998) Circular permutation of beta B2-crystallin changes the hierarchy of domain assembly. *Protein Sci.* **7**, 1280–1285
14. Smith, M. A., Bateman, O. A., Jaenicke, R., and Slingsby, C. (2007) Mutation of interfaces in domain-swapped human beta B2-crystallin. *Protein Sci.* **16**, 615–625
15. Vendra, V. P. R., Agarwal, G., Chandani, S., Talla, V., Srinivasan, N., and Balasubramanian, D. (2013) Structural integrity of the greek key motif in beta gamma-crystallins is vital for central eye lens transparency. *PLoS One* **8**, e70336
16. Mackay, D. S., Boskovska, O. B., Knopf, H. L. S., Lampi, K. J., and Shiels, A. (2002) A nonsense mutation in CRYBB1 associated with autosomal dominant cataract linked to human chromosome 22q. *Am. J. Hum. Genet.* **71**, 1216–1221
17. Yang, J. H., Zhu, Y. H., Gu, F., He, X., Cao, Z. F., Li, X. X., *et al.* (2008) A novel nonsense mutation in CRYBB1 associated with autosomal dominant congenital cataract. *Mol. Vis.* **14**, 727–732
18. Deng, H., and Yuan, L. (2014) Molecular genetics of congenital nuclear cataract. *Eur. J. Med. Genet.* **57**, 113–122
19. Wang, K. J., Wang, S., Cao, N. Q., Yan, Y. B., and Zhu, S. Q. (2011) A novel mutation in CRYBB1 associated with congenital cataract-microcornea syndrome: the p.Ser129Arg mutation destabilizes the betaB1/betaA3-crystallin heteromer but not the betaB1-crystallin homomer. *Hum. Mutat.* **32**, E2050–E2060
20. Xi, Y. B., Zhao, W. J., Zuo, X. T., Tjondro, H. C., Li, J., Dai, A. B., *et al.* (2014) Cataract-causing mutation R233H affects the stabilities of betaB1- and betaA3/betaB1-crystallins with different pH-dependence. *Biochim. Biophys. Acta* **1842**, 2216–2229
21. Jin, A., Zhang, Y., Xiao, D., Xiang, M., Jin, K., and Zeng, M. (2019) A novel mutation p.S93R in CRYBB1 associated with dominant congenital cataract and microphthalmia. *Curr. Eye Res.* **45**, 483–489
22. Leng, X. Y., Li, H. Y., Wang, J., Qi, L. B., Xi, Y. B., and Yan, Y. B. (2016) Congenital microcornea-cataract syndrome-causing mutation X253R increases betaB1-crystallin hydrophobicity to promote aggregate formation. *Biochem. J.* **473**, 2087–2096
23. Qi, L. B., Hu, L. D., Liu, H., Li, H. Y., Leng, X. Y., and Yan, Y. B. (2016) Cataract-causing mutation S228P promotes betaB1-crystallin aggregation and degradation by separating two interacting loops in C-terminal domain. *Protein Cell* **7**, 501–515
24. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858
25. Willoughby, C. E., Shafiq, A., Ferrini, W., Chan, L. L. Y., Billingsley, G., Priston, M., *et al.* (2005) CRYBB1 mutation associated with congenital cataract and microcornea. *Mol. Vis.* **11**, 587–593
26. Wang, J., Ma, X., Gu, F., Liu, N. P., Hao, X. L., Wang, K. J., *et al.* (2007) A missense mutation S228P in the CRYBB1 gene causes autosomal dominant congenital cataract. *Chin. Med. J. (Engl.)* **120**, 820–824
27. Shiels, A., and Hejtmancik, J. F. (2017) Mutations and mechanisms in congenital and age-related cataracts. *Exp. Eye Res.* **156**, 95–102
28. Shiels, A., and Hejtmancik, J. F. (2019) Biology of inherited cataracts and opportunities for treatment. *Annu. Rev. Vis. Sci.* **5**, 123–149
29. Shiels, A., and Hejtmancik, J. F. (2021) Inherited cataracts: genetic mechanisms and pathways new and old. *Exp. Eye Res.* **209**, 108662
30. Xian, C. X., Zhu, M. W., Nong, T. Y., Li, Y. Q., Xie, X. M., Li, X., *et al.* (2021) A novel mutation in ext2 caused hereditary multiple exostoses through reducing the synthesis of heparan sulfate. *Genet. Mol. Biol.* **44**, e20200334
31. Chiu, J., March, P. E., Lee, R., and Tillett, D. (2004) Site-directed, ligase-independent mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h. *Nucleic Acids Res.* **32**, e174
32. Lu, G., and Gong, P. (2013) Crystal structure of the full-length Japanese encephalitis virus NS5 reveals a conserved methyltransferase-polymerase interface. *PLoS Pathog.* **9**, e1003549

33. Wang, Q. S., Yu, F., Huang, S., Sun, B., Zhang, K. H., Liu, K., *et al.* (2015) The macromolecular crystallography beamline of SSRF. *Nucl. Sci. Tech.* **26**, 12–17
34. Pflugrath, J. W. (1999) The finer things in X-ray diffraction data collection. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 1718–1725
35. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
36. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
37. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., *et al.* (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954
38. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921
39. Theobald, D. L., and Wuttke, D. S. (2006) THESEUS: maximum likelihood superpositioning and analysis of macromolecular structures. *Bioinformatics* **22**, 2171–2172