Published in final edited form as:

J Struct Biol. 2012 December ; 180(3): 439–446. doi:10.1016/j.jsb.2012.07.011.

Crystal structure of the globular domain of C1QTNF5: Implications for late-onset retinal macular degeneration

Xiongying Tu¹ and Krzysztof Palczewski^{1,*}

¹Department of Pharmacology, School of Medicine, Case Western Reserve University

Abstract

Autosomal dominant late-onset retinal macular degeneration (L-ORMD) is caused by a single S163R mutation in the C1q and tumor necrosis factor-related protein 5 (*C1QTNF5*) gene. The *C1QTNF5* gene encodes a secreted and membrane-associated protein involved in adhesion of retinal pigmented epithelial cells (RPE) to Bruch's membrane. The crystal structure of the trimeric globular domain of human C1QTNF5 at 1.34 Å resolution reveals unique features of this novel C1q family member. It lacks a Ca²⁺-binding site, displays a remarkable non-uniform distribution of surface electrostatic potentials and possesses a unique sequence (F₁₈₁F₁₈₂G₁₈₃G₁₈₄W₁₈₅P₁₈₆) that forms a hydrophobic plateau surrounded by Lys and Arg residues with a solvent cavity underneath. S₁₆₃ forms a hydrogen bond with F₁₈₂ in a hydrophobic area extending to the hydrophobic plateau. The pathogenic mutation S163R disrupts this hydrogen bonding and positively charges these hydrophobic areas. Thus, our analysis provides insights into the structural basis of the L-ORMD disease mechanism.

Keywords

L-ORMD; L-ORD; age-related macular degeneration; AMD; CTRP5; drusen

1. Introduction

Late-onset retinal macular degeneration (L-ORMD) is a fully penetrant autosomal dominant disorder characterized by bilateral loss of vision, abnormal dark-adaption, drusen-like yellow spots in the fundus, choroidal neovascularization, retinal atrophy and long anteriorly inserted lens zonules (Borooah et al., 2009; Jacobson et al., 2001; Kuntz et al., 1996; Milam et al., 2000). L-ORMD is caused by a single S163R missense mutation in the globular C1q domain of the C1q and tumor necrosis factor-related protein 5 (*C1QTNF5*) (Hayward et al., 2003). Affected patients consistently exhibit thick extracellular deposits of lipid-rich material between the retinal pigmented epithelium (RPE) and Bruch's membrane, which eventually lead to RPE dysfunction, photoreceptor cell death and consequent loss of sight. Major clinical features of L-ORMD are similar to those seen in age-related macular degeneration (AMD), a complex disorder affecting up to 30% of the elderly population and accounting for 50–60% of new blind registration in developed countries (Green and Enger, 1993; Hayward et al., 2003). L-ORMD thus can be used to model disease mechanisms of

© 2012 Elsevier Inc. All rights reserved.

*Correspondence to: Krzysztof Palczewski, Ph.D., Department of Pharmacology, School of Medicine, Case Western Reserve University, 10900 Euclid Ave, Cleveland, Ohio 44106–4965, USA; Phone: 216–368–4631; Fax: 216–368–1300; kxp65@case.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

AMD by virtue of its monogenic etiology. Moreover, studies of L-ORMD may lead to the discovery of treatments applicable to both disorders. Other members of this family when mutated cause diseases such as systemic lupus erythematosus, insulin-resistant diabetes, Schmid's metaphyseal chondrodysplasia and others (Ghebrehiwet et al., 2012).

The *C1QTNF5* gene encodes a 25 kDa secreted and membrane-associated protein highly and specifically expressed in the RPE and ciliary body of the eye and adipose tissue (Hayward et al., 2003; Mandal et al., 2006; Wong et al., 2008). C1QTNF5 contains three domains: a signal peptide domain (residues 1–15), a collagen domain (residues 30–98) composed of 23 uninterrupted Gly-X-Y repeats, and a gC1q domain (residues 103–243) (Fig. 1A). C1QTNF5 is a novel member of C1q family that shares a common feature of trimerization (Shapiro and Scherer, 1998) in which three protomers are intertwined into a 'globular head' (trimeric gC1q domain) situated at the C-terminus of a collagen 'stalk' (collagen triple helix). Although studies of complement protein C1q have shown that these formed trimers further multimerize into a bouquet-like superstructure through their interacting collagen domains (Fig. 1B), direct evidence is still needed to clarify that C1QTNF5 can also adopt this quaternary structure. Current studies suggest that N-terminal cysteine residues (C₂₈ and C₉₈) are involved in higher order multimerisation of C1QTNF5 trimers via disulphide bonds (Wong et al., 2008).

The disease mechanism(s) induced by the S163R mutation in C1QTNF5 is not well understood. The S163R mutant was shown to be unstable and prone to aggregation (Hayward et al., 2003). Moreover, the S163R mutant protein causes a marked loss of cellular adhesion (Shu et al., 2006a). A recently generated heterozygous knock-in mouse model carrying the disease-associated mutation largely recapitulates the pathological features of L-ORMD patients (Chavali et al., 2011). We here report structural studies of the human C1QTNF5 globular head. This structure provides insights into the structural mechanism of the pathogenic S163R mutation, reveals certain unique features of C1QTNF5, discloses the role of conserved hydrophobic residues in the folding and assembly of C1q family proteins, and explains how certain mutagenic changes lead to the disruption of trimerization.

2. Materials and Methods

2.1. Protein Expression and Purification

A codon-optimized, full-length human C1QTNF5 gene, with natural NcoI restriction site eliminated, was synthesized and cloned into a pUC57 vector. The C-terminal gC1q domain (nucleotides 306–729) of this construct was amplified by PCR. The forward PCR primer introduced an NcoI restriction site at the 5'-end, and the reverse primer introduced codons for six His residues and an XhoI restriction site at the 3'-end. The purified PCR product and expression vector pETDuet-1 were digested with restriction enzymes NcoI and XhoI, and then ligated together in a molar ratio of 3 to 1. Verified by DNA sequencing, the recombinant plasmid construct was transformed into RosettaTM(DE3)pLysS Competent cells (Novagen). An overnight culture from a single colony was inoculated into fresh liquid lysogeny broth (LB) in a volume ratio of 1 to 50 and grown to an OD_{600 nm} of 0.6 at 37 °C. The cell culture was subsequently cooled to 16 °C and induced overnight with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation, resuspended in buffer A (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1 mM β-mercaptoethanol, and 10% glycerol) supplemented with protease inhibitors (cOmplete EDTA-free, Roche), disrupted on ice by sonication, and cleared by centrifugation at 4 °C. The supernatant was applied to 4 ml of Ni-NTA affinity resin in a 1.5 cm × 6.3 cm column through gravity chromatography. The Ni-NTA column was extensively washed with buffer B (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 1 mM β-mercaptoethanol, 10% glycerol, and 25 mM

imidazole), and then eluted with buffer C (Buffer B with the imidazole concentration elevated to 300 mM). Eluted protein was further purified by gel filtration chromatography on a Superdex 200 10/300 GL column (GE Healthcare). Peak fractions containing C1QTNF5 were pooled and concentrated to 2 mg/ml with 10 k Amicon[®] ultracentrifugal filters. The purified protein migrated on a precast 4–15% sodium dodecyl sulfate polyacrylamide gel (Bio-Rad) as a single band with an approximate molecular mass 16 kDa, which corresponded to the expected molecular weight of the C1QTNF5 protomer head. The apparent molecular weight of native C1QTNF5 was about 45 kDa based on the elution profile from the gel filtration.

2.2. Crystallization and Diffraction Data Collection

Crystallization was performed at room temperature by a traditional hanging-drop method. Two μ l of protein sample was mixed with an equal volume of reservoir solution (0.1 M sodium acetate pH 5.0, 2 M NaCl, and 10% glycerol), and equilibrated against 500 μ l of reservoir solution. Crystals grew to sizes suitable for X-ray diffraction in two days. Crystals were soaked in cryoprotectant solution (reservoir solution with the glycerol concentration elevated to 25%) for ~30 s, and then immediately flash-cooled in liquid nitrogen. Diffraction data were collected at the beamline 24-ID-C of the Advanced Photon Source.

2.3. Structural Determination and Analysis

Diffraction data were processed, reduced and scaled with software HKL2000 (Otwinowski and Minor, 1997). Initial phases were derived from the structure of mouse adiponectin (PDB ID: 1C28) (Shapiro and Scherer, 1998) by the molecular replacement method implemented in Phaser (McCoy et al., 2007) from the CCP4 software package (Winn et al., 2011) with the following statistics: 7.6 of rotation function Z-score (RFZ), 9.6 of translation function Z-score (TFZ) and 142 of log-likelihood gain (LLG). After the first round of simulated annealing refinement and composite omit map calculation, the starting model was trimmed to its core structure by deleting regions that apparently did not fit into the electron density maps. The starting model (protomer A of the complement protein C1q globular head) shared only 39% sequence identity with its counterpart in C1QTNF5, so *in silico* mutations, deletions and insertions were carried out to build a complete model. Iterative model building and refinement were conducted using Coot (Emsley and Cowtan, 2004) and Phenix (Adams et al., 2010), respectively. Anisotropic B factor was refined at the end of refinement. Composite omit maps and simulated annealing omit maps were calculated with CNS (Brunger et al., 1998). The final model and structural factor have been deposited in the Protein Data Bank (www.pdb.org) under accession code 4F3J. The quality of the final model was analyzed by PROCHECK (Laskowski et al., 1993). Structural alignment was achieved with secondary-structure matching (Krissinel and Henrick, 2004) implemented in the Coot software. Surface electrostatic potentials were calculated by APBS implemented in the PyMOL software (Baker et al., 2001; Schrödinger). *In silico* mutagenesis of G149A and S163R was achieved with Coot, and the side-chain torsional angles of R163 were chosen by two criteria, namely minimization of steric clashes with neighboring residues and the use of side-chain torsional angles preferred by other C1q family proteins at this site. Figures were made with software PyMOL and Jalview (Waterhouse et al., 2009).

3. Results

3.1. Overall structure of the human C1QTNF5 trimeric head

The structure of the C1QTNF5 trimeric head was determined to a resolution of 1.34 Å in the *H3* space group with an R_{work} of 13.6% and an R_{free} of 16.5% (Table 1). Electron density maps allowed us unambiguously to trace the entire main chain and model most of the side chains. Biochemical studies have shown that the globular head of C1QTNF5 is a

homotrimer (Shu et al., 2006b). The structure shows that C1QTNF5 globular head has the same mode of trimerization as other members of the C1q family. Each individual protomer adopts a 10-strand jelly-roll fold arranged in two anti-parallel 5-stranded β -sheets (A', A, H, C, F) and (E, D, G, B, B') (Fig. 1C and S1), with nomenclature which conforms to that of the tumor necrosis factor (TNF) family (Jones et al., 1989). The two β -sheets are folded between strand E and strand F, forming two 5-strand layers with an inner layer (A', A, H, C, F) buried inside. There are some conserved hydrophobic residues including F₁₀₆, L₁₂₃, F₁₂₅, Y₁₃₆, F₁₄₃, C₁₄₅, Y₁₅₁, F₁₅₃, V₁₉₇, L₁₉₉, V₂₀₅, V₂₀₇, F₂₂₇ and L₂₃₁, which are widely scattered throughout their β -strands (Fig. 1C). Pinpointing these residues on the structure of a protomer of C1QTNF5 reveals that they cluster into a hydrophobic 'zipper' that brings the two β -sheet layers together (Fig. 2B), suggesting a concerted role of these conserved residues in folding of the tertiary structure. The homotrimer is formed by packing the inner layers together. Residues accounting for this trimerization are mostly located in strands H, C, F and part of E. These residues are almost entirely hydrophobic near the base, and become progressively more hydrophilic toward the top of the trimer, forming a solvent cavity at the central interface (Fig. 2C). Y₁₅₂, F₂₃₀ and V₂₃₂ near the base form a hydrophobic core that is critical for trimer formation. Sequence alignment shows that these residues are highly conserved throughout the C1q family (Fig. 1C). Our structure reveals that the hydrophobic core resides adjacent to the hydrophobic 'zipper' box (Fig. 2B). An aromatic motif, identified by comparing the sequences of collagens X, VIII, complement factor C1q and fibrillar collagens (Brass et al., 1992), is also found in C1QTNF5 (residues 143–153 in human C1QTNF5, Fig. 1C). This aromatic motif is at the border of the 'zipper' box and hydrophobic core and accommodates both a key residue for trimerization (Y₁₅₂) and critical residues for forming the 'zipper' box (F₁₄₃, C₁₄₅, Y₁₅₁, F₁₅₃, V₁₉₇, L₁₉₉, and L₂₃₁). Structural rearrangement of this 'zipper' box thus would alter its ability to mediate trimer formation. Unlike conserved hydrophobic residues forming the 'zipper', the conserved hydrophilic residues form hydrogen bonds within their 5-strand layer (Supplementary Table 1).

3.2. Structural comparison of gC1q domains from C1QTNF5, complement protein C1q and adiponectin

C1q family proteins have evolved diverse functions but maintain a similar structural scaffold (Ghebrehewet et al., 2012). Although the sequence identity of gC1q domains among members of this family is lower than 40%, their structural superposition showed overall root-mean-square deviation (rmsd) values of ~ 0.5 Å for β -strands and ~ 1.8 Å for loops for C α atoms, respectively (Supplementary Fig. 1). Loops and β -strands often serve different roles in protein structure and function. Loops underlie the versatile functions of C1q family proteins, as exemplified by the well-studied complement protein C1q. Its loops A|A' and B'|B of protomer C recognize "eat me" signals on apoptotic cell surfaces (Paidassi et al., 2008), but the same loops of protomer B recognize IgG (Gaboriaud et al., 2003). The β -strands contribute more significantly to the folding and assembly of C1q family proteins. A typical example is strand C and H, both of which are involved in trimerization.

Despite an overall conserved structural scaffold, C1QTNF5 displays unique features that are not found in the other family members. C1QTN5 lacks Ca²⁺-binding sites that however are essential to structural stability and functions of other C1q family members. Adiponectin and collagen X NC1 possess three Ca²⁺-binding sites at identical positions in their respective structures (Min et al., 2012). Their essential coordination residues (D₁₈₇ and D₁₉₅ for adiponectin; D₆₂₆ and D₆₃₄ for collagen X NC1) are conserved throughout their classes and aligned at equivalent sites (Fig. 1C). Complement protein C1q contains only one Ca²⁺-binding site because of its heterotrimeric arrangement (Gaboriaud et al., 2003), the essential coordination residue being D₁₇₂ of protomer B (Fig. 1C). However, C1QTNF5 lacks all of

these conserved essential coordination residues at its equivalent sites. The electron density maps were carefully examined where Ca^{2+} -binding sites occur in other C1q family proteins, but no potential electron densities of Ca^{2+} ions were observed, even though the trimeric head was crystallized in the presence of 10 mM CaCl_2 .

3.3. Solvent cavity at the central interface

A discontinuous solvent cavity was found along the central interface of the homotrimeric head of C1QTNF5. This solvent cavity has a perfect 3-fold symmetry, in sharp contrast with its asymmetric counterpart in complement protein C1q. Moreover, this cavity is characterized by a series of elaborate hydrophobic and hydrophilic interactions (Fig. 2C). The solvent cavity originating at the base is blocked by the aforementioned hydrophobic core at one third of its height from the base of the structure, forming a cone-shape solvent cavity with its large circumference at the bottom. The hydroxyl group of Y₁₅₂ in the hydrophobic core interacts with a water molecule located 3.04 Å away at the 3-fold axis and forms a hydrogen bond with the imidazole group of the conserved H₁₅₆ from another protomer 2.94 Å distant (Fig. 2C). The imidazole group of H₁₅₆ further interacts with the hydroxyl group of the conserved S₁₉₂ at a distance of 2.80 Å. Starting from the location of Y₁₅₂, the wall of this solvent cavity becomes increasingly hydrophilic, and is finally closed by three hydrophobic W₁₈₅ residues, one from each protomer at the apex of the solvent cavity (Fig. 2C and Supplementary Fig. 2). Solvent cavities of other members of C1q family with known structures are not sealed by hydrophobic residues but are freely accessible to external solvent instead.

3.4. Surface properties of the C1QTNF5 trimeric head

The gC1q domain of human C1QTNF5 consists of 60% non-polar residues, only 9.2% of negatively charged residues (13 Asp and Glu) and 7.8% of positively charged residues (11 Arg and Lys), but its surface exhibits an intriguing electrostatic potential pattern (Fig. 3a). On top of the C1QTNF5 head there is a predominantly hydrophobic triangular plateau with an area of 108 Å² composed of F₁₈₁, W₁₈₅ and P₁₈₆ residues (Fig. 2D and Supplementary Fig. 2). Extending downward from each F₁₈₁ edge is a hydrophobic area where S₁₆₃, the disease-associated mutation site, is buried. This hydrophobic area comprising about half the size of the plateau, is composed of F₁₇₉, F₁₈₂, A₁₆₂ and Y₂₁₄. In light of the ability of C1QTNF5 to associate with the RPE plasma membranes (Mandal et al., 2006), it should be emphasized that this plateau is the only continuous hydrophobic area on the C1QTNF5 head surface. Four sets of positively charged spikes extend outward from the surface just below the plateau in a step-wise manner. These charged spikes are formed by side chains of K₁₈₇, R₁₆₁, K₂₂₂ and R₁₁₄ of each protomer. Moreover, these four residues enclose a positively charged cavity visible from the top surface (Fig. 3a, top view). The central bottom surface displays a markedly positive charge potential compared to other members of this protein family, which marks the solvent cavity and is due primarily to R₁₀₃ (Fig. 3a, bottom view). The side surface of the C1QTNF5 head, however, is predominantly negatively charged (Fig. 3a, side view). This study reveals a significantly non-uniform charge distribution that contrasts sharply with that of complement protein C1q and adiponectin. On the top surface of the complement protein C1q head there are positively charged spikes composed of side chains of K₁₇₃ from protomer A and K₁₇₀ from protomer C that guard the entrance to the central solvent cavity with its inner surface distinctly filled with negative charges (Fig. 3b, top view). In contrast to both complement protein C1q and C1QTNF5, the top surface of the adiponectin globular head is dominated by a negative charge potential (Fig. 3c, top view). Moreover, as opposed to the negative-charge predominance on the side surface of C1QTNF5, both complement protein C1q and adiponectin feature an overall positive charge potential on their side surfaces (Fig. 3, side view). Both bottom surfaces of human

complement protein C1q and adiponectin reveal a predominantly positive charge, but its extent is far less remarkable than that of C1QTNF5 (Fig. 3, bottom view).

3.5. The disease-associated S₁₆₃ to R₁₆₃ mutation

S₁₆₃ resides at an apex loop formed by strand C and D (Fig. 4A). This residue is buried within a hydrophobic area mainly composed of A₁₆₂, F₁₇₉, F₁₈₁, F₁₈₂, and Y₂₁₄ (Fig. 4). Surprisingly, these residues just suffice to enwrap S₁₆₃, mutation of which (S163R) causes L-ORMD. The structure also reveals that the hydroxyl group of S₁₆₃ interacts with the main-chain nitrogen of F₁₈₂ through hydrogen bonding at a distance of 2.89 Å (Fig. 4A). F₁₈₂ is also the second residue of the apical loop (residues 181–188, Fig. 4A) which forms the previously mentioned hydrophobic triangular plateau with a solvent cavity underneath. Hydrogen bonding stabilizes and positions the apical loop from each protomer and the disease-causing S163R mutation abolishes this interaction (Fig. 4B). Moreover, the guanidinium group of R₁₆₃ also disrupts the hydrophobic area in which it resides. This potential effect of S163R is supported by our modeling study showing that R₁₆₃ significantly alters this uncharged area, transforming it to a positively charged one (Fig. 4C and D).

4. Discussion

The members of the C1q family are highly diversified with respect to their amino acid sequences, functions and tissue localization, but share a similar tertiary and quaternary structure. This structural conservation despite marked sequence variability prompted us to investigate the determinants of the folding and assembly of their tertiary and quaternary structures. Through protein sequence alignment within the C1q family, we identified some conserved sites which mostly consist of hydrophobic residues. These hydrophobic residues are widely scattered, but our structural analysis revealed that they actually cluster into a strip, resembling a zipper that brings two 5-strand β sheets together. Deletion of the conserved hydrophobic 143–153 motif in the C1QTNF5 head has been shown to result in a complete loss of trimerization (Shu et al., 2006a). The importance of this motif for trimerization is clearly explained by its location at the border of the ‘zipper’ box and the hydrophobic core, which accommodates residues from both (Fig. 2). Thus, both F143A and F153A ‘zipper’ box mutations have been found to abolish trimerization of the C1QTNF5 head (Shu et al., 2006a). These two residues are located exactly at the conserved aromatic motif. Another mutation in the same motif, G149A, also causes disruption of trimerization (Shu et al., 2006a). G₁₄₉ resides at β -strand C of the structure (Fig. 1C) and neighbors residues 231 and 233. When changed to Ala residue in our structure, the side chain sterically clashes with the side chain of L₂₃₁ and the main-chain oxygen of Y₂₃₃ at distances of 2.67 Å and 2.28 Å, respectively (Supplementary Fig. 3). These unfavorable clashes likely result in a local structural displacement, which can spread to neighboring Y₁₅₁, F₂₃₀ and V₂₃₂ residues. L₂₃₁ itself and Y₁₅₁ are components of the ‘zipper’ box, whereas F₂₃₀ and V₂₃₂ are components of the hydrophobic core, thus this could alter the conformations of both the zipper box and the hydrophobic box and abolish trimerization of C1QTNF5. Based on our structure, we deduce that the hydrophobic ‘zipper’ determines the folding of its tertiary structure and thereby also affects trimerization. This study also reveals how the C1q family maintains its conserved structural scaffold while evolving diverse functions.

Some unique features of C1QTNF5 are its distinct surface electrostatic properties and the entrance of a solvent cavity at its apex. An apical loop forms the entrance to the solvent cavity on top of the central interface among members of the C1q family. The main-chain conformation of the C1QTNF5 apical loop, however, is profoundly displaced from that seen in adiponectin, collagen X NC1, and complement protein C1q. A characteristic hydrophobic segment (F₁₈₁F₁₈₂G₁₈₃G₁₈₄W₁₈₅P₁₈₆) is found within this apical loop and is only unique to

C1QTNF5. The arrangement of the C1QTNF5 apical loop is caused by hydrophobic-stacking of F₁₈₁ and W₁₈₅ (Fig. 2D and Supplementary Fig. 2). The three hydrophobic segments, one from each protomer, shape the apex into a hydrophobic triangular plateau with the three W₁₈₅ residues inside to gate the underlying solvent cavity along the central interface (Supplementary Fig. 2). In contrast to C1QTNF5, apex regions of complement protein C1q, adiponectin and collagen X NC1 are hydrophilic and their underlying cavities are directly accessible to external solvent. The hydrophobic plateau surrounded by significant positively-charged spikes on the downward side is another significant feature that distinguishes C1QTNF5 from other members of the C1q family. Considering that C1QTNF5 can associate with RPE plasma membranes, it is very likely that the top region of C1QTNF5 is responsible for this property. This would explain the structural basis by which bouquet-like C1QTNF5 can enhance adhesion of RPE cells to themselves or Bruch's membrane. Studies has shown that the S163R mutant is unstable and prone to aggregation (Hayward et al., 2003), and results in a marked loss of cellular adhesion (Shu et al., 2006a). One hypothesis about the disease mechanism is that S163R mutant has an effect on reducing the adhesion of RPE cells (Hayward et al., 2003; Shu et al., 2006a), which in turn increase probability of built-up of lipid-rich material between RPE cells and Bruch's membrane in affected aging patients. This also explains why this disease is late-onset. The structure of the C1QTNF5 head shows that S₁₆₃ forms a hydrogen bond with F₁₈₂, and is buried within a hydrophobic region extending to the hydrophobic plateau. Our study suggests that the S163R mutation eliminates this hydrogen bond and significantly increases the positive electrostatic potential of these two hydrophobic areas (Fig. 4B and D). The S163R mutation therefore likely causes displacement of the apical loop and an expansion of its surrounding hydrophobic areas and those on the plateau, which will result in aggregation of S163R mutant and, as a consequence, reduce the adhesion of RPE cells. Notably, regions around residues equivalent to S₁₆₃ of C1QTNF5 are all hydrophilic in other family proteins with known structures. Another hypothesis about the disease mechanism is that the S163R mutation causes misfolding of this protein (Hayward et al., 2003; Shu et al., 2006a). This however appears unlikely because this mutation is located in the apex loops (the apical loop and loop of S₁₆₃) with tertiary and quaternary structures determined by the 'zipper' and hydrophobic core that are intact on the opposite side. It is also notable that this site is not conserved and can also be occupied by an Asp residue in adiponectin, an Asn residue in collagen VIII NC1 and complement protein C1q and a His residue in collagen X NC1. These observations indicate that S₁₆₃ is not the crucial residue for protein folding.

Interestingly, examination of C1q family structures reveals that all these residues share a hydrogen bond similar to that of S₁₆₃ in C1QTNF5. Conservation of this hydrogen bonding suggests an important functional role for the apical loop. The only exception is protomer A of complement protein C1q, which has a Glu residue at this site instead. The longer side chain of E₁₆₃ causes potential interacting groups to locate further away from each other, resulting in a loss of this critical hydrogen bond in the structure of protomer A. This observation further indicates that the loss of this hydrogen bond should not lead to misfolding of C1QTNF5. The heterotrimeric property of complement protein C1q may be the reason that it can afford the loss of this particular hydrogen bond in one protomer.

In conclusion, our structure indicates alterations of the hydrophobic apex region by the S163R mutation result in reduced cell adhesion and increased aggregation. These findings underlie the critical functional role of the C1QTNF5 apex region. Considering the signature feature of C1QTNF5 stated above, it is the apex region that binds RPE plasma membranes and enhances RPE cell adhesion. The dominant pathological effect of the S163R mutation is likely caused by the loss-of-function when the mutant protein coexists with wild-type protein in heterozygous individuals. Aggregation of the S163R mutant in the homozygous

state primarily results from rearrangement of the hydrophobic apex region that in turn causes exposure of hydrophobic side chains.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Drs. Philip D. Kiser and David L. Lodowski for help with data collection. We also thank Drs. Leslie T. Webster Jr., David L. Lodowski, and members of Palczewski's laboratory (Case Western Reserve University) for valuable comments on the manuscript. This research was supported in part by grants EY008061 and P30EY11373 from the National Institutes of Health. The work also is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines supported by grants from the National Center for Research Resources (5P41RR015301-10) and the National Institute of General Medical Sciences (8P41GM103403-10) from the National Institutes of Health. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357.

References

- Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography*. 2010; 66:213–221.
- Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America*. 2001; 98:10037–10041. [PubMed: 11517324]
- Borooah S, Collins C, Wright A, Dhillon B. Republished review. [corrected] Late-onset retinal macular degeneration: clinical insights into an inherited retinal degeneration. *Postgraduate medical journal*. 2009; 85:495–500. [PubMed: 19734518]
- Brass A, Kadler KE, Thomas JT, Grant ME, Boot-Handford RP. The fibrillar collagens, collagen VIII, collagen X and the C1q complement proteins share a similar domain in their Cterminal non-collagenous regions. *FEBS letters*. 1992; 303:126–128. [PubMed: 1607009]
- Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta crystallographica. Section D, Biological crystallography*. 1998; 54:905–921.
- Chavali VR, Khan NW, Cukras CA, Bartsch DU, Jablonski MM, et al. A CTRP5 gene S163R mutation knock-in mouse model for late-onset retinal degeneration. *Human molecular genetics*. 2011; 20:2000–2014. [PubMed: 21349921]
- Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography*. 2004; 60:2126–2132.
- Francis K, van Beek J, Canova C, Neal JW, Gasque P. Innate immunity and brain inflammation: the key role of complement. *Expert reviews in molecular medicine*. 2003; 5:1–19. [PubMed: 14585169]
- Gaboriaud C, Juanhuix J, Gruez A, Lacroix M, Darnault C, et al. The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. *The Journal of biological chemistry*. 2003; 278:46974–46982. [PubMed: 12960167]
- Ghebrehiwet B, Hosszu KK, Valentino A, Peerschke EI. The C1q family of proteins: insights into the emerging non-traditional functions. *Frontiers in immunology*. 2012; 3
- Green WR, Enger C. Age-related macular degeneration histopathologic studies. The 1992 Lorenz E. Zimmerman Lecture. *Ophthalmology*. 1993; 100:1519–1535. [PubMed: 7692366]
- Hayward C, Shu X, Cideciyan AV, Lennon A, Barran P, et al. Mutation in a short-chain collagen gene, CTRP5, results in extracellular deposit formation in late-onset retinal degeneration: a genetic model for age-related macular degeneration. *Human molecular genetics*. 2003; 12:2657–2667. [PubMed: 12944416]

- Jacobson SG, Cideciyan AV, Wright E, Wright AF. Phenotypic marker for early disease detection in dominant late-onset retinal degeneration. *Investigative ophthalmology & visual science*. 2001; 42:1882–1890. [PubMed: 11431457]
- Jones EY, Stuart DI, Walker NP. Structure of tumour necrosis factor. *Nature*. 1989; 338:225–228. [PubMed: 2922050]
- Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta crystallographica. Section D, Biological crystallography*. 2004; 60:2256–2268.
- Kuntz CA, Jacobson SG, Cideciyan AV, Li ZY, Stone EM, et al. Sub-retinal pigment epithelial deposits in a dominant late-onset retinal degeneration. *Investigative ophthalmology & visual science*. 1996; 37:1772–1782. [PubMed: 8759344]
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography*. 1993; 26:283–291.
- Mandal MN, Vasireddy V, Reddy GB, Wang X, Moroi SE, et al. CTRP5 is a membrane-associated and secretory protein in the RPE and ciliary body and the S163R mutation of CTRP5 impairs its secretion. *Investigative ophthalmology & visual science*. 2006; 47:5505–5513. [PubMed: 17122142]
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. Phaser crystallographic software. *Journal of applied crystallography*. 2007; 40:658–674. [PubMed: 19461840]
- Milam AH, Curcio CA, Cideciyan AV, Saxena S, John SK, et al. Dominant late-onset retinal degeneration with regional variation of sub-retinal pigment epithelium deposits, retinal function, and photoreceptor degeneration. *Ophthalmology*. 2000; 107:2256–2266. [PubMed: 11097607]
- Min X, Lemon B, Tang J, Liu Q, Zhang R, et al. Crystal structure of a single-chain trimer of human adiponectin globular domain. *FEBS letters*. 2012; 586:912–917. [PubMed: 22449980]
- Otwinowski Z, Minor W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology*. 1997:307–326.
- Paidassi H, Tacnet-Delorme P, Garlatti V, Darnault C, Ghebrehiwet B, et al. C1q binds phosphatidylserine and likely acts as a multiligand-bridging molecule in apoptotic cell recognition. *Journal of immunology*. 2008; 180:2329–2338.
- Schrödinger L. The PyMOL Molecular Graphics System. Version 1.5.0.1.
- Shapiro L, Scherer PE. The crystal structure of a complement-1q family protein suggests an evolutionary link to tumor necrosis factor. *Current biology: CB*. 1998; 8:335–338. [PubMed: 9512423]
- Shu X, Tulloch B, Lennon A, Vlachantoni D, Zhou X, et al. Disease mechanisms in late-onset retinal macular degeneration associated with mutation in C1QTNF5. *Human molecular genetics*. 2006a; 15:1680–1689. [PubMed: 16600989]
- Shu X, Tulloch B, Lennon A, Hayward C, O'Connell M, et al. Biochemical characterisation of the C1QTNF5 gene associated with late-onset retinal degeneration. A genetic model of age-related macular degeneration. *Advances in experimental medicine and biology*. 2006b; 572:41–48. [PubMed: 17249553]
- Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. 2009; 25:1189–1191. [PubMed: 19151095]
- Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography*. 2011; 67:235–242.
- Wong GW, Krawczyk SA, Kitidis-Mitrokostas C, Revett T, Gimeno R, et al. Molecular, biochemical and functional characterizations of C1q/TNF family members: adipose-tissue-selective expression patterns, regulation by PPAR-gamma agonist, cysteine-mediated oligomerizations, combinatorial associations and metabolic functions. *The Biochemical journal*. 2008; 416:161–177. [PubMed: 18783346]

Fig. 1.

Domain organization of C1QTNF5 and sequence alignment of members of the C1q family proteins. **A.** The domain structure of C1QTNF5. **B.** A bouquet-like arrangement of the C1q family proteins. Figure adapted from (Francis et al., 2003) with permission from the authors and publisher. The three protomers are shown in magenta, green and cyan. The trimeric globular head is outlined by a black box. The top and base of the globular head are indicated by arrows. **C.** Sequence alignment of gC1q domains from C1q family proteins. Sequences are those of human C1QTNF5, human complement protein C1q (PDB ID: 1PK6), human collagen X NC1 (PDB ID: 1GR3), human collagen VIII NC1 (PDB ID: 1O91), and human adiponectin (PDB ID: 4DOU). Conserved residues are highlighted in white letters with background colors ranging from blue to black according to their conservation. The conserved hydrophobic motif (residues 143–153) and the F₁₈₁F₁₈₂G₁₈₃G₁₈₄W₁₈₅P₁₈₆ sequence are outlined in red. S₁₆₃ is marked with an asterisk. Key aspartic acids coordinating with Ca²⁺ ions are highlighted in red. Secondary structure is assigned according to the structure of the trimeric head of human C1QTNF5.

Fig. 2. Structure of the homotrimeric head of human C1QTNF5. **A.** Overall structure. **B.** Closeup view of the hydrophobic 'zipper' of protomer A and the hydrophobic core of the central interface. **C.** Close-up view of the solvent cavity at the central interface (protomer B is not shown). **D.** Close-up view of the hydrophobic plateau on the top of the trimeric head. Protomers A, B and C are colored gray, pink, and yellow, respectively. The vertical lines showed in the overall structure represent 3-fold axis at the central interface. Carbon atoms of the conserved hydrophobic residues that form the 'zipper' are shown in cyan for protomer A. The 'zipper' box and hydrophobic core are outlined with red and black ovals, respectively. Water molecules are represented by orange spheres. Hydrogen bond distances have unit of Å.

Fig. 3.

Comparison of electrostatic potentials on surfaces of selected structures of known C1q family proteins viewed from the top, side and bottom. Molecular surfaces are colored according to their electrostatic potentials with a color scale set from -7 kT/e (red) to 7 kT/e (blue), as calculated by the Adaptive Poisson-Boltzmann Solve (APBS) implemented in PyMOL software (Baker et al., 2001; Schrödinger). Positive potentials are shown in blue, negative potentials in red, and hydrophobicity in white. (a) The homotrimeric head of human C1QTNF5; (b) the heterotrimeric head of human complement protein C1q; (c) the homotrimeric head of human adiponectin.

Fig. 4.

Pathogenic S163R mutation. **A.** S₁₆₃ forms a hydrogen bond with F₁₈₂. The hydrogen bond of S₁₆₃ positions and stabilizes the apical loop. S₁₆₃ is wrapped by hydrophobic residues A₁₆₂, F₁₇₉, F₁₈₁, F₁₈₂, and Y₂₁₄. **B.** Arg163 disrupts this hydrogen bond, changing the electrostatic potentials in this area and causing separation of the involved residues. Electrostatic potentials of surfaces around S₁₆₃ (**C**) and modeled R₁₆₃ (**D**) of human C1QTNF5. The molecular surfaces are colored according to electrostatic potential with color scale set from -7 kT/e (red) to 7 kT/e (blue), as calculated by APBS implemented in PyMOL. Positive potentials are shown in blue, negative potentials in red, and hydrophobicity is in white. Hydrogen bond distances have unit of Å.

Table 1

Statistics of Data Collection and Refinement

Data collection	
Space group	<i>H3</i>
Unit cell dimensions	
a = b (Å)	46.62
c (Å)	138.83
Resolution range (Å)	46.28 – 1.34 (1.39–1.34)
Unique reflections	122,037
R_{merge}^b	0.066 (0.324)
$\langle I \rangle / \langle \sigma I \rangle^c$	20.4 (3.0)
Redundancy	4.8 (2.5)
Completeness (%)	98.3 (89.1)
Refinement Statistics	
Resolution range (Å)	46.28–1.34
Number of atoms	
Protein	1,116
Water	90
Rms deviation	
Bond length (Å)	0.009
Bond angle (°)	1.373
$R_{\text{work}}/R_{\text{free}}^d$ (%)	13.6/16.5
Ramachandran analysis (%) ^e	
Most favored	85.7
Additionally allowed	12.6
Generously allowed	1.7
Disallowed	0.0

^aSelected statistics from data collection and refinement are shown. Values in parentheses are data from the highest resolution shell.

^b $R_{\text{merge}} = \sum |hkl| \sum_j |I_j - \langle I \rangle| / \sum |hkl| \sum_j I_j$. $\langle I \rangle$ is the mean intensity of j observations of reflection hkl and its symmetry equivalents.

^c $\langle I \rangle / \langle \sigma I \rangle$ is the mean intensity divided by the mean error.

^d R_{work} is $\sum |F_o| - |F_c| / \sum |F_o|$ where F_o is an observed amplitude and F_c is a calculated amplitude; R_{free} is the same statistic calculated over a subset of the data that was not used for refinement.

^eRamachandran analysis from PROCHECK.