

MUTATION IN BRIEF

De Novo Gene Conversion in the RCA Gene Cluster (1q32) Causes Mutations in Complement Factor H Associated With Atypical Hemolytic Uremic Syndrome

Stefan Heinen², Pilar Sanchez-Corral⁷, Michael S Jackson¹, Lisa Strain¹, Judith A Goodship¹, Elizabeth J Kemp¹, Christine Skerka², T Sakari Jokiranta³, Kevin Meyers⁴, Eric Wagner⁵, Pierre Robitaille⁵, Jorge Esparza-Gordillo⁶, Santiago Rodriguez de Cordoba⁶, Peter F Zipfel², and Timothy H J Goodship^{1*}

¹The Institute of Human Genetics, University of Newcastle upon Tyne, Newcastle upon Tyne, United Kingdom; ²Department for Infection Biology, Hans Knoell Institute for Natural Products Research, Jena, Germany; ³Department of Bacteriology and Immunology, Haartman Institute and Helsinki University Central Hospital, University of Helsinki, Helsinki, Finland; ⁴The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; ⁵Ste. Justine Hospital, Montreal, Québec, Canada; ⁶Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Madrid, Spain; ⁷Unidad de Investigación, Hospital Universitario La Paz, Madrid, Spain

*Correspondence to: Professor Tim Goodship, Institute of Human Genetics, International Centre for Life, Central Parkway, Newcastle upon Tyne, NE1 3BZ, United Kingdom; E-mail: t.h.j.goodship@ncl.ac.uk

Grant sponsor: Kidney Research UK, the Northern Counties Kidney Research Fund, the Foundation for Children with atypical HUS and the Robin Davies Trust.

Communicated by Mark H. Paalman

Many of the complement regulatory genes within the RCA cluster (1q32) have arisen through genomic duplication and the resulting high degree of sequence identity is likely to predispose to gene conversion events. The highest degree of identity is between the genes for factor H (CFH) and five factor H-related proteins – CFHL1, CFHL2, CFHL3, CFHL4, and CFHL5. CFH mutations are associated with atypical hemolytic uremic syndrome (aHUS). In the Newcastle cohort of 157 aHUS patients we have identified CFH mutations in 25 families or individuals. Eleven of these 25 independent mutations are either c.3226C>G,Q1076E; c.3572C>T,S1191L; c.3590T>C,V1197A or combined c.3572C>T,S1191L/c.3590T>C,V1197A. Sequence analysis shows that all four of these changes could have arisen as a result of gene conversion between CFH and CFHL1. Analysis of parental samples in two patients with S1191L/V1197A has shown that the changes are de novo thus providing conclusive evidence that gene conversion is the mutational mechanism in these two cases. To confirm that S1191L and V1197A are disease predisposing we examined their functional significance in three ways – analysis of the C3b/C3d binding characteristics of recombinant mutant S1191L/V1197A protein, heparin affinity chromatography and haemolytic assays of serum samples from aHUS patients carrying these changes. The results showed that these changes resulted in impaired C3b

Received 15 August 2005; accepted revised manuscript 20 November 2005.

binding and a defective capacity to control complement activation on cellular surfaces. We, therefore, provide conclusive evidence that gene conversion is responsible for functionally significant CFH mutations in aHUS. © 2006 Wiley-Liss, Inc.

KEY WORDS: HUS; complement; factor H; gene conversion

INTRODUCTION

The RCA (Regulators of Complement Activation) gene cluster on human chromosome 1q32 spans a total of 21.45 cM and contains more than 60 genes of which 15 are complement related genes. These code for both soluble and transmembrane proteins which play a pivotal role in regulating complement activity. These complement genes are arranged in tandem within two groups. In a centromeric 360 kb segment lie the genes for factor H (*CFH*) (MIM# 134370) and five factor H-related proteins – *CFHL1* (MIM# 134371), *CFHL2* (MIM# 600889), *CFHL3* (MIM# 605336), *CFHL4* (MIM# 605337) and *CFHL5* (MIM# 608593) (aliases *FHRI-5*) (Fig. 1).

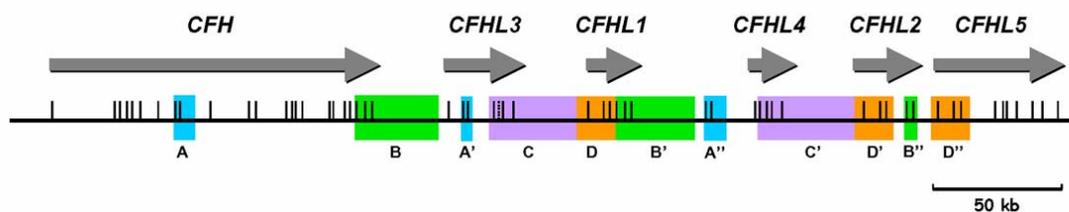


Figure 1. Position of the genes for CFH and the factor H related proteins in a centromeric segment of the RCA cluster at 1q32. Genomic duplications including the different exons of the 6 genes are shown by (A-D) and colour coding. Exons are indicated as vertical lines

Sequence analysis of this region shows evidence of several large genomic duplications resulting in a high degree of sequence identity between *CFH* and the genes for the five factor H-related proteins (Zipfel, et al., 1999). The secreted protein products of these genes are similar in that they consist of repetitive units (~60 amino acids) named short consensus repeats (SCR) or complement control protein modules (CCP). Each SCR is generally encoded by a single exon. *CFH* consists of 20 SCRs and *CFHL1* five. The highest degree of sequence identity is seen between SCRs 18-20 of *CFH* and SCRs 3-5 of *CFHL1* (Appendix 1). *CFH* SCR 18 and *CFHL1* SCR 3 consist of 59 amino acids. At the nucleotide level the exons encoding these two SCRs differ by 5 bases. 3 result in an amino acid difference. *CFH* has tyrosine, valine and glutamine residues at positions 1058, 1060 and 1076 respectively (encoded by triplets TAT, GTG and CAA) whilst *CFHL1* has histidine, leucine, and glutamic acid residues at positions 157, 159 and 175 (encoded by triplets CAT, CTG and GAA). *CFH* SCR 19 and *CFHL1* SCR 4 consist of 61 amino acids. The exons encoding these two SCRs differ at one nucleotide position which does not result in a coding change. *CFH* SCR20 and *CFHL1* SCR 5 consist of 67 amino acids. The exons differ at two nucleotide positions, both of which affect the amino acid sequence of the encoded proteins; *CFH* has serine and valine residues at positions 1191 and 1197 (encoded by triplets TCG and GTT) while *CFHL1* has leucine and alanine residues at positions 290 and 296 (encoded by triplets TTG and GCT).

CFH mutations are associated with atypical hemolytic uremic syndrome (aHUS; MIM# 235400) (Richards, et al., 2001), a disease characterised by the triad of thrombocytopenia, Coomb's test negative microangiopathic haemolytic anemia and acute renal failure (Moake, 2002). Mutations reported include c.3572C>T, S1191L and c.3590T>C, V1197A either singly or in combination (Caprioli, et al., 2001; Perez-Caballero, et al., 2001; Richards, et al., 2001) raising the possibility that gene conversion of *CFH* SCR 20 by SCR 5 of *CFHL1* is the mutational mechanism in a proportion of cases. c.3226C>G, Q1076E has also been reported (Neumann, et al., 2003; Richards, et al., 2001) raising the possibility of gene conversion of *CFH* SCR 18 by SCR 3 of *CFHL1*. We report here 2 aHUS cases where S1191L and V1197A changes have occurred in combination as de novo events, thus providing unambiguous evidence that gene conversion is the mutational mechanism involved. In addition we report a series of aHUS patients in whom Q1076E, S1191L, and V1197A have been found either singly or in combination

providing further support for our hypothesis. Finally, we provide evidence that these changes are functionally significant.

MATERIALS AND METHODS

Newcastle cohort of aHUS patients

This consists of 157 aHUS patients for whom clinical details and biological samples including DNA, serum and plasma are available. Within this cohort there are 26 families where more than one affected family member is included. The study was approved by the Northern and Yorkshire Multi-Centre Research Ethics Committee and all subjects gave informed consent. In this cohort, we have identified *CFH* mutations in 25 aHUS families or individuals. Of these 25 independent mutations, 3 (including the 2 aforementioned de-novo cases) are S1191L and V1197A together, 4 are S1191L (including two families of three sibs and two sibs respectively), 2 are V1197A and 2 are Q1076E. In total 14 patients (7 male, 7 female), some of whom have been reported previously (Donne, et al., 2002; Olie, et al., 2004; Olie, et al., 2005; Richards, et al., 2001), had these changes. Their age at first presentation ranged from 11 months to 30 years and all developed end stage renal failure. These 4 changes represent 11/25 (44%) of *CFH* mutations in our cohort.

Patients with de novo combined S1191L and V1197A

Two patients, one male and one female presented with aHUS at the age of 11 months. Both developed end stage renal failure requiring dialysis. One has since received a cadaver renal transplant and one year post transplant remains free of either recurrent disease or rejection.

DNA Analysis

Genomic DNA was extracted from peripheral blood by standard methods and then amplified by PCR using exon specific primers and annealing temperatures as reported previously (Richards, et al., 2001). PCR products were purified using magnetic microparticles (AMPure, Agencourt Biosciences) to remove unincorporated dNTPs, primers and salts. Sequencing reactions were carried out by dye terminator cycle sequencing using either exon specific primers or M13 primers, purified using magnetic microparticles (CleanSeq, Agencourt Biosciences) and then electrophoresed on a fluorescent 16 capillary sequencer (Beckman CEQ 8000). The resulting electropherograms were checked for homozygous and heterozygous base changes using an automated sequence analysis package (Mutation SurveyorTM 5.0, <http://www.softgenetics.com>). The reference nucleotide sequence for *CFH* is taken from GenBank RefSeq-file NM_000186.1 and the nucleotide numbering uses the A of the ATG translation initiation start site as nucleotide +1. The factor H amino acid numbering includes the 18-residue signal peptide. The reference nucleotide sequence for *CFHL1* is taken from GenBank RefSeq-file NM_002113.1 and the nucleotide numbering uses the A of the ATG translation initiation start site as nucleotide +1. The factor H related protein 1 amino acid numbering includes the 18-residue signal peptide.

To exclude the possibility that an uncharacterised mutation allowed annealing of the *CFH* SCR20 specific primers to the *CFHL1* gene (causing amplification of CFHL1 SCR5 sequence), we also generated sequence using *CFH* specific primers which encompassed 3 upstream single nucleotide differences (in the terminal introns of *CFH* and *CFHL1*), and 4 downstream single nucleotide differences (in the 3'UTR of the two genes) in a single amplicon (Fig. 2).

Functional assessment of S1191L and V1197A

To confirm that S1191L and V1197A are disease predisposing changes we examined their functional significance in three ways – analysis of the C3b/C3d binding characteristics of recombinant mutant S1191L/V1197A protein, heparin affinity chromatography and haemolytic assays of serum samples from aHUS patients carrying these changes. Recombinant fragments representing the C-terminal domain of CFH (CFH18-20) and a fragment that has mutations at positions S1191L and V1197A were expressed and assayed for C3b and C3d binding by surface plasmon resonance as described previously (Manuelian, et al., 2003). C3b or C3d (CalBiochem) were coupled by standard amine-coupling procedure in equal amounts to a flow cell on a sensor chip (carboxylated dextran chip CM5, BIAcore AB). Identical concentrations (2 µg/ml) of ligand were used. Heparin affinity chromatography was

undertaken on serum from S1191L, V1197A and S1191L/V1197A patients (one of each) (Manuelian, et al., 2003). The individual fractions were separated by SDS-PAGE and assayed by Western blotting. Haemolytic assays were undertaken as described previously (Sanchez-Corral, et al., 2004) on serum from patients carrying S1191L/V1197A, S1191L and V1197A mutations. These assays use sheep erythrocytes and are performed in the presence of Mg-EGTA to prevent activation of the classical pathway of complement.

RESULTS

DNA analysis

DNA analysis of parental samples showed that in two of the three patients with combined S1191L/V1197A the changes were de novo. Parental sample integrity for both patients was confirmed by the PowerPlex[®] 16 System (Promega). For each of the de-novo cases the only changes from the reference *CFH* sequence were in the 2 coding nucleotides in the exon encoding CFH SCR 20 (Fig. 2) confirming the specificity and localised nature of the mutational events. To confirm that the changes in these two patients occurred on the same allele subcloned PCR products were sequenced.

```

CTGATGTAATGAATCATTGATAATACACCCCTAATTCTCATAACATTAAACATCGAACCTCA
CTGATGTAATGAATCATTGATAATACACCCCTAATTCTCATAACATTAAACATAAAACCTCA

TTTTCACATCGATTACCATTCTAAGTTTATTCAAATCAATATGATGTTTTCTACATAGTTGG
TTTTCACATCGATTACCATTTAAGTTTATTTAAATCAATATGATGTTTTTACATAGTCGG

TTTGGATAGTGTTTTGAGAAATAATTCCTGAACCATCATATAACATTCTACTTGAAAACCT
TTTGGACAGTGTTTTGAGAAATAATTCCTGAACCATCATATAACATTCTACTTGAAAACCT

GAAAGTCTATGAAGATTTGCATACTACTTAATGTTTTATGTTTACTGTTTTTTATTTTCAG
GAAAGTCTATGAAGATTTGCATACTACTTAATGTTTTATGTTTACTGTTTTTTATTTTCAG

ATCCGTGTGTAATATCCCGAGAAATTATGGAAAATTATAACATAGCATTAAAGGTGGACAGC
ATCCGTGTGTAATATCCCGAGAAATTATGGAAAATTATAACATAGCATTAAAGGTGGACAGC

CAAACAGAAGCTTTATTGAGAACAGGTGAATCAGTTGAATTTGTGTGTAAACGGGGATAT
CAAACAGAAGCTTTATTTGAGAACAGGTGAATCAGCTGAATTTGTGTGTAAACGGGGATAT

CGTCTTTTCATCACGTTTCTCACACATTGCGAACCAACATGTTGGGATGGGAAACTGGAGTATC
CGTCTTTTCATCACGTTTCTCACACATTGCGAACCAACATGTTGGGATGGGAAACTGGAGTATC

CAACTTGTGCAAAAAGATAGAATCAATCATAAAGTGCACACCTTTATTTCAGAACTTTAGTA
CAACTTGTGCAAAAAGATAGAATCAATCATAAAAATGCACACCTTTATTTCAGAACTTTAGTA

TTAAATCAGTTCTCAATTTTCAATTTTTATGTATTGTTTTACTCCTTTTTTATTTCATACGTAA
TTAAATCAGTTCTTAATTTTCAATTTTTAA-GTATTGTTTTACTCCTTTTTTATTTCATACGTAA

AATTTTGGATTAATTTGTGAAAATGTAATTATAAGCTGAGACCGGTGGCTCTCTTCTTAAA
AATTTTGGATTAATTTGTGAAAATGTAATTATAAGCTGAGACCGGTGGCTCTCTTCTTAAA

AGCACCATATTAATACTCTGGAAAATAACGGTTGTGTC
AGCACCATATTAATACTTGGAAAATAACTGTTGTGTC

```

Figure 2. CFH (above) and CFHL1 (below) sequence for the exons (in red) encoding CFH SCR 20 and CFHL1 SCR 5. Intronic and 3' UTR sequence is in black. Sequence differences are highlighted. The respective primer pairs are indicated by single and double underlining.

Functional studies

Analysis of the C3b/C3d binding characteristics showed that wild type CFH₁₈₋₂₀ bound to immobilized C3b and C3d, as indicated by the prominent association and dissociation profile (Fig. 3). In contrast the profile for mutant CFH₁₈₋₂₀ suggests that it binds with lower affinity to both C3b and C3d (Fig. 3).

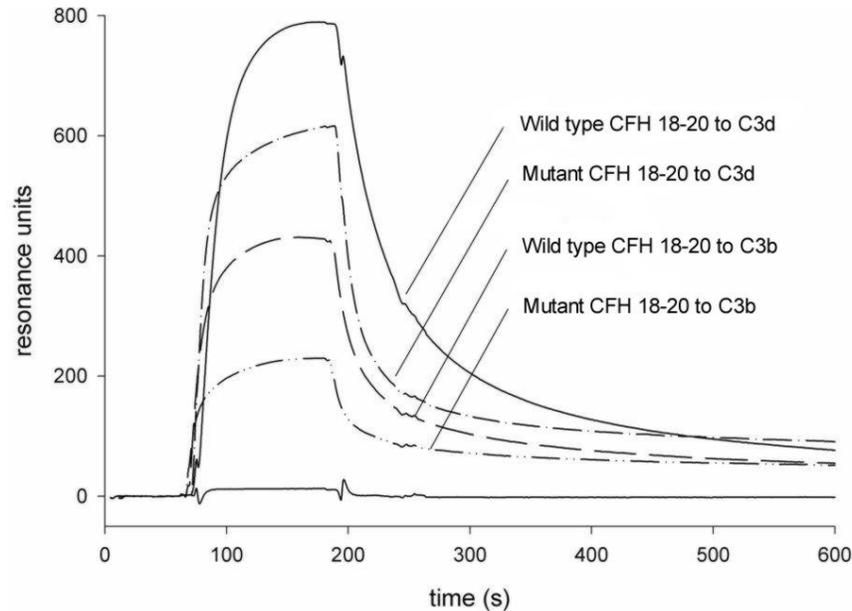


Figure 3. Binding of wild type and mutant (S1191L and V1197A) CFH SCR 18-20 to C3b and C3d analysed by surface plasmon resonance. The wild type and mutant proteins were injected at a concentration of 2 $\mu\text{g}/\text{ml}$ into a flow cell coupled with either C3b or C3d. The binding of the mutant to both C3b and C3d is decreased. All recombinant proteins were tested on two different chip surfaces and representative figures are shown.

Heparin Affinity Chromatography (reviewed but not shown) did not show a significant difference in heparin binding of the mutant proteins. Haemolytic assays showed that in all patient samples the capacity to control complement activation on cellular surfaces was defective (Fig. 4). That this defect was present in heterozygous patients suggests that the wild type protein present was not able to compensate for the mutant protein.

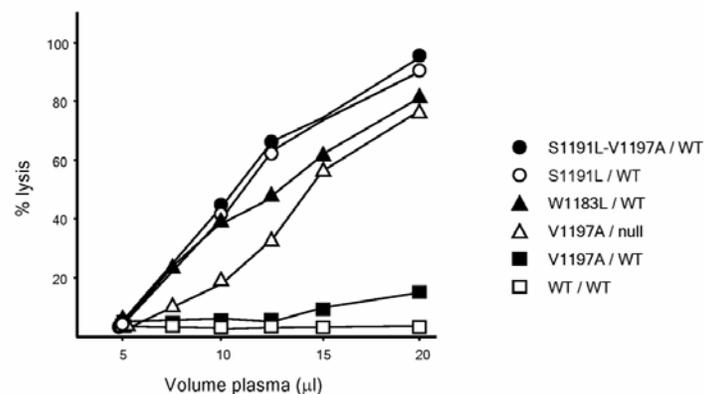


Figure 4. Lysis of sheep erythrocytes by serum from three aHUS patients with heterozygous S1191L, V1197A, or S1191L/V1197A and one patient with hemizygous V1183L. The result from a patient with a different mutation W1183L is shown for comparison. Lysis is shown as a function of the volume of serum added and is represented as a percentage of the control of total lysis. In these assays, lysis of sheep erythrocytes reveals an anomalous regulation of the alternative pathway of complement activation in the tested serum.

DISCUSSION

In this study we have shown that gene conversion of CFH SCR20 by CFHL1 SCR5 is the only plausible mutational mechanism which can account for the observed de novo double mutations, with the flanking sequence variation defining the conversion tract length as being between 19 and 331 nucleotides in length. For these changes to have occurred through nucleotide substitution we would be forced to hypothesise two independent germline mutations occurring in the same gene in a single generation in two unrelated individuals at precisely the same positions. Our functional analyses show that the mutations at position 1191 and 1197 affect the binding characteristics of CFH protein to two cleavage products of the central complement component C3. CFH SCRs 18-20 are known to be important for binding to both C3b and anionic substances (Rodriguez de Cordoba, et al., 2004). The functional data presented here lead us to hypothesise that it is primarily C3b binding that is affected by S1191L and V1197A. Mutant CFH would, therefore, bind to host cell surfaces but its ability to bind C3b and thus regulate complement activation is impaired.

Therefore, we provide conclusive evidence that gene conversion is responsible for functionally significant de novo *CFH* mutations in two cases of aHUS. Similar mutations are seen in over a third of aHUS cases in the Newcastle series known to have a *CFH* mutation and it is probable that they too have arisen by gene conversion. Whilst circumstantial evidence for gene conversion in other diseases such as spinal muscular atrophy and hereditary pancreatitis is available (Chen, et al., 2000; Frugier, et al., 2002) reports of de-novo changes such as seen in congenital adrenal hyperplasia (Collier, et al., 1993) are extremely rare.

REFERENCES

- Caprioli J, Bettinaglio P, Zipfel PF, Amadei B, Daina E, Gamba S, Skerka C, Marziliano N, Remuzzi G, Noris M. 2001. The molecular basis of familial hemolytic uremic syndrome: Mutation analysis of factor H gene reveals a hot spot in short consensus repeat 20. *Journal of the American Society of Nephrology* 12:297-307.
- Chen JM, Ragueneas O, Ferec C, Deprez PH, Verellen-Dumoulin C. 2000. A CGC>CAT gene conversion-like event resulting in the R122H mutation in the cationic trypsinogen gene and its implication in the genotyping of pancreatitis. *Journal of Medical Genetics* 37:E36.
- Collier S, Tassabehji M, Sinnott P, Strachan T. 1993. A de novo pathological point mutation at the 21-hydroxylase locus: implications for gene conversion in the human genome. *Nature Genetics* 3:260-5.
- Donne RL, Abbs I, Barany P, Elinder CG, Little M, Conlon P, Goodship THJ. 2002. Recurrence of hemolytic uremic syndrome after live related renal transplantation associated with subsequent de novo disease in the donor. *American Journal of Kidney Diseases* 40:E22.
- Frugier T, Nicole S, Cifuentes-Diaz C, Melki J. 2002. The molecular bases of spinal muscular atrophy. *Current Opinion in Genetics and Development* 12:294-8.
- Manuelian T, Hellwage J, Meri S, Caprioli J, Noris M, Heinen S, Jozsi M, Neumann HP, Remuzzi G, Zipfel PF. 2003. Mutations in factor H reduce binding affinity to C3b and heparin and surface attachment to endothelial cells in hemolytic uremic syndrome. *Journal of Clinical Investigation* 111:1181-1190.
- Moake JL. 2002. Thrombotic microangiopathies. *New England Journal of Medicine* 347:589-600.
- Neumann HP, Salzmann M, Bohnert-Iwan B, Manuelian T, Skerka C, Lenk D, Bender BU, Cybulla M, Riegler P, Konigsrainer A and others. 2003. Haemolytic uraemic syndrome and mutations of the factor H gene: a registry-based study of German speaking countries. *Journal of Medical Genetics* 40:676-681.
- Olie KH, Florquin S, Groothoff JW, Verlaak R, Strain L, Goodship TH, Weening JJ, Davin JC. 2004. Atypical relapse of hemolytic uremic syndrome after transplantation. *Pediatric Nephrology* 19:1173-1176.
- Olie KH, Goodship THJ, Verlaak R, Florquin S, Groothoff JW, Strain L, Weening JJ, Davin JC. 2005. Post transplantation cytomegalovirus-induced recurrence of atypical haemolytic uremic syndrome associated with a factor H mutation: successful treatment with intensive plasma exchanges and ganciclovir. *American Journal of Kidney Diseases* 45:e12.

- Perez-Caballero D, Gonzalez-Rubio C, Gallardo ME, Vera M, Lopez-Trascasa M, Rodriguez de Cordoba S, Sanchez-Corral P. 2001. Clustering of missense mutations in the C-terminal region of factor H in atypical hemolytic uremic syndrome. *American Journal of Human Genetics* 68:478-84.
- Richards A, Buddles MR, Donne RL, Kaplan BS, Kirk E, Venning MC, Tielemans CL, Goodship JA, Goodship THJ. 2001. Factor H mutations in hemolytic uremic syndrome cluster in exons 18-20, a domain important for host cell recognition. *American Journal of Human Genetics* 68:485-490.
- Rodriguez de Cordoba S, Esparza-Gordillo J, Goicoechea de Jorge E, Lopez-Trascasa M, Sanchez-Corral P. 2004. The human complement factor H: functional roles, genetic variations and disease associations. *Molecular Immunology* 41:355-67.
- Sanchez-Corral P, Gonzalez-Rubio C, Rodriguez de Cordoba S, Lopez-Trascasa M. 2004. Functional analysis in serum from atypical hemolytic uremic syndrome patients reveals impaired protection of host cells associated with mutations in factor H. *Molecular Immunology* 41:81-4.
- Zipfel PF, Jokiranta TS, Hellwage J, Koistinen V, Meri S. 1999. The factor H protein family. *Immunopharmacology* 42:53-60.

Appendix 1. CFH and CFHL1 sequence for the exons encoding CFH SCR3-5 and CFHL1 SCR3-5.

