Vol. 194, No. 1, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 15, 1993 Pages 458-464

POLYMORPHISM OF HUMAN COMPLEMENT COMPONENT C6: AN AMINO ACID SUBSTITUTION (GLU/ALA) WITHIN THE SECOND THROMBOSPONDIN REPEAT DIFFERENTIATES BETWEEN THE TWO COMMON ALLOTYPES C6 A AND C6 B

Georg Dewald, Markus M. Nöthen, and Sven Cichon

Institute of Human Genetics, University of Bonn, Wilhelmstr. 31, 53111 Bonn, Germany

Received May 27, 1993

Summary. Component C6 of the human complement system exhibits a genetic polymorphism in all populations tested so far. Using isoelectric focusing two common allotypes, C6 A ('acidic') and C6 B ('basic') and a number of rare variants have been described. A comparison of the two published cDNA sequences of C6 suggests a polymorphism in codon 98. Using polymerase chain reaction (PCR) we amplified a segment of the human C6 gene encompassing the presumably polymorphic codon. According to the restriction fragment patterns obtained after DdeI digestion of the PCR product, three genotypes were distinguished. The polymorphism is caused by a nucleotide substitution (C→A) in the second position of codon 98; allele 1 (GCG) codes for Ala, allele 2 (GAG) for Glu. Sequencing of PCR products confirmed the mutation. For 46 unrelated individuals genotyped by this PCR-based method we also determined C6 protein phenotype. Three phenotypes were observed (C6 A, C6 AB, C6 B). There was an absolute concordance between C6 protein typing and DNA typing. We thus conclude that the C6 A allotype is characterized by a Glu and the C6 B allotype by an Ala residue in position 98 of the C6 polypeptide chain.

Complement component C6 is one of the five hydrophilic precursor proteins of the membrane attack complex (MAC) of the complement system. The MAC is a high-molecular amphiphilic heteropolymer which is assembled from the terminal complement components C5, C6, C7, C8, and C9 upon proteolytic activation of C5 (1,2). The membranolytic activity of MAC is of importance for host defence against invading microorganisms. The deposition of sublytic amounts of MAC on target membranes can stimulate, at least in some nucleated cell types, cellular activities like the production of inflammatory mediators (3).

A genetically determined polymorphism of human C6 was first described by Hobart, Lachmann and Alper (4,5). The preferred method for the characterization of C6 phenotypes is isoelectric focusing in polyacrylamide gels (4,6); the detection of C6 is performed either by a specific hemolytic overlay (4,5,7) or an immunoblotting procedure (8,9). C6 variants are coded by codominant alleles at an autosomal locus. There are two very common alleles, *C6*A* ('acidic') and *C6*B* ('basic'), and at least 19 rare structural variants (6). Genetic deficiency of C6, a condition associated with an increased risk for meningococcal infections (10), appears to be due to a null allele (*C6*Q0*) at the C6 locus (11). In Caucasians the two common alleles have frequencies of about 0.61 (*C6*A*) and 0.38 (*C6*B*) (4, 7); in Mongoloids the *C6*B*

0006-291X/93 $4.00

frequency ($\approx$0.50) is higher than the *C6*A* frequency ($\approx$0.44) (9,12), whereas for Negroids allele frequencies of 0.55 (*C6*A*) and 0.40 (*C6*B*) have been reported (13). These numbers mean that there is a remarkably high heterozygosity rate in all the major racial groups.

The C6 locus is tightly linked to the locus coding for complement component C7 (14-16) and also to the gene coding for component C9 (16,17). All three genes have been assigned to chromosome 5 (18,19). These data suggest the existence of a gene cluster that includes three structurally and functionally related terminal complement proteins.

Human C6 is a single-chain plasma glycoprotein with a calculated molecular weight of 108,000 (20). The complete primary structure of C6 was derived independently by two groups from the sequence of C6 cDNA clones (20,21). The polypeptide chain consists of 913 amino acids. A comparison of the two published amino acid sequences suggests the presence of either a glutamic acid residue or an alanine residue at position 98, corresponding to an A→C nucleotide substitution in the second position of codon 98. We postulated that such a type of amino acid exchange (negatively charged → neutral) could explain the different isoelectric point of C6 A and C6 B. Using polymerase chain reaction we have developed a simple method for direct investigation of the nucleotide sequence in codon 98; this method allowed us to demonstrate that codon 98 is indeed polymorphic and to examine the relationship between this DNA polymorphism and the C6 protein polymorphism (22).

Material and Methods

Blood samples. EDTA anticoagulated venous blood samples were collected from 46 unrelated German individuals. Plasma was obtained by centrifugation and leucocyte DNA was isolated as described by Miller et al. (23). In addition, DNA samples from four two-generation families were used for inheritance studies.

Polymerase chain reaction (PCR). The PCR was performed using the following primers: C6-4-F (5'-TCTGTCTTGCGTCCCAGTC-3') corresponding to positions 468 - 486 of the C6 cDNA sequence reported by Haefliger et al. (21) and C6-4-R (5'-TCTTGCAGTCAGCCTCTTCA-3') complementary to positions 579 - 560. The primers were synthesized on an Applied Biosystems DNA Synthesizer PCR-Mate. Standard PCR was carried out in a 50 μl volume containing 80 ng genomic template, 20 pmol of each primer, 200 μM of each deoxynucleotide triphosphate (dATP, dCTP, dGTP, dTTP), 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.01% gelatine and 2.5 U Taq-Polymerase (Perkin Elmer Cetus). Samples were overlaid with mineral oil and were processed in a Perkin-Elmer Cetus DNA Thermal Cycler through 35 temperature cycles consisting of 30 sec at 94°C, 30 sec at 55°C, and 30 sec at 72°C. After the last cycle the samples were incubated at 72°C for 10 min to allow the final extension step to be completed. An aliquot (25 μl) of the PCR product was digested using 20 U/sample of DdeI (New England Bio Labs) according to the manufacturer's instructions. Fragments were separated in 10% polyacrylamide gels and visualized by ethidium bromide staining. Fragment sizes were estimated by comparison with a 10 base pair ladder (Life Technologies).

Cloning and sequencing of PCR products. PCR products from two individuals were cloned in pCR-Script SK(+) vector (Stratagene). A hemibiotinylated PCR product was generated by using one biotinylated vector primer and one normal vector primer. The PCR product was incubated with streptavidine Dynabeads (Dynal Ltd.) and magnetic beads were collected with a magnetic concentrator. After washing and denaturing the bound strand of DNA was sequenced by the dideoxy nucleotide chain termination method (24) using Sequenase Version 2.0 Kit (US Biochemicals).

Analysis of C6 protein polymorphism. C6 allotypes were determined by methods previously described (9,25), with some minor modifications. Plasma samples were subjected to isoelectric

focusing (5,100 volthours) in thin-layer polyacrylamide gels (145 x 240 x 0.8 mm) containing 2.2% (w/v) Ampholine carrier ampholytes [Pharmacia; pH 5-8 (6 parts) and pH 3.5-10 (1 part)] and 0.2 M taurine. The focused proteins were transferred onto a nitrocellulose membrane (Schleicher & Schuell, BA-85) by press-blotting for 1 h (25). After blocking in phosphate-buffered saline containing 3% bovine serum albumin (Sigma) C6 patterns were detected by means of an enzyme immunoassay using goat anti-human-C6 antiserum (Cappel; diluted 1:300) and peroxidase-conjugated rabbit anti-goat-immunoglobulin antiserum (DAKO; diluted 1:600). Peroxidase activity was developed using o-dianisidine. C6 phenotypes were scored without knowledge of genotyping results.

Results

The two published C6 cDNA sequences suggest a sequence difference in codon 98. To evaluate if there is indeed a sequence polymorphism in the population we used PCR to amplify a 112-base pair (bp) segment of the C6 gene, containing the putatively polymorphic region. Primers were designed from the published C6 cDNA sequence (21,20) assuming homology of the C6 gene structure with the known exon-intron organization of complement C9 gene (26). Amplification of genomic DNA resulted in a single DNA fragment of the expected size (Fig. 1).

The sequence difference noted in the cDNA sequences is distinguishable by a change in a recognition site for restriction endonuclease DdeI. As can be seen from Fig. 1 digestion of the 112-bp PCR product with DdeI resulted in the expected restriction fragment length polymorphism (RFLP). The presence of a constant DdeI restriction site in the amplified fragment provides a convenient internal digestion control. Three genotypes (11, 12, 22) were observed each displaying a unique combination of DdeI fragment sizes. Cleavage in the constant (non-polymorphic) DdeI site produced in all individuals a fragment of size 35 bp. Then, depending on the absence or presence of the polymorphic DdeI restriction site, either a fragment of 77 bp (allele 1) or two fragments of 41 bp and 36 bp (allele 2) were produced (Fig. 1). The existence of a DNA sequence variation in codon 98 was confirmed by sequencing the PCR product from

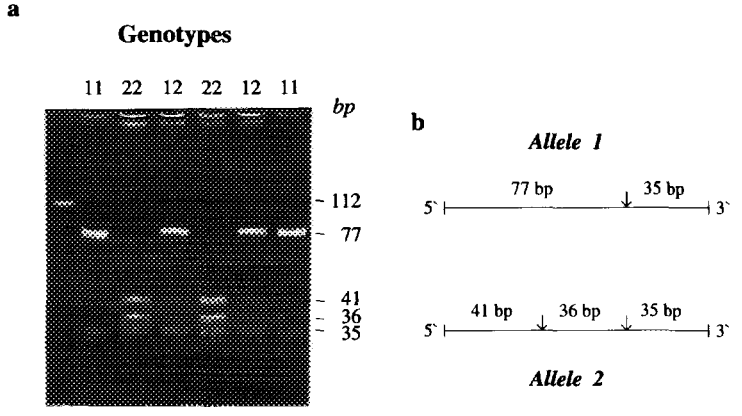

Figure 1. Restriction fragment length polymorphism of complement C6 gene as detected by DdeI digestion of a 112-bp PCR product encompassing codons 84-121. **a.** Electrophoretic separation of PCR products; lane 1: undigested PCR product; lanes 2-7: DdeI digested PCR products of six different individuals representing the heterozygous and the two homozygous genotypes. **b.** Schematic diagram of the restriction fragments obtained by DdeI digestion of the PCR product. Arrows indicate DdeI recognition sites.

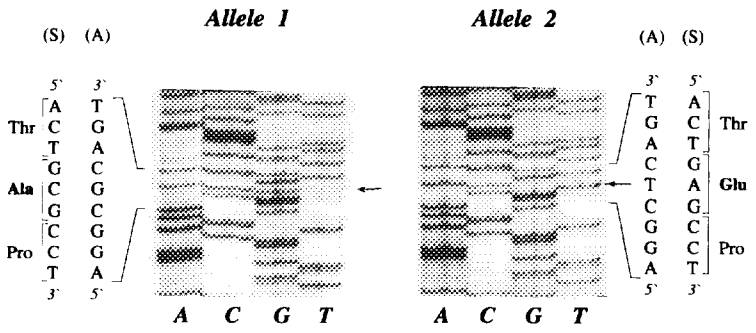

Figure 2. Autoradiograms of the regions of sequencing gels showing the point mutation that causes polymorphism of the C6 gene. *Left*: Sequence analysis of amplified DNA from an individual with genotype 11 (and protein phenotype C6 B); *right*: Sequence analysis of amplified DNA from an individual with genotype 22 (and protein phenotype C6 A). (S) and (A) are sense- and antisense-strand sequence, respectively. Arrows indicate the single nucleotide substitution that determines the presence of either Glu or Ala at position 98 of the mature C6 protein.

two genotypically homozygous individuals [subject K121 (genotype 11) and subject K123 (genotype 22)] (Fig. 2). Typing 46 unrelated Caucasians we obtained the following allele frequencies: 0.38 (allele 1; codon 98: GCG) and 0.62 (allele 2; codon 98: GAG). Autosomal codominant inheritance of alleles was demonstrated in four two-generation families each with at least two children.

To investigate the relationship between the DdeI RFLP and C6 protein polymorphism plasma samples from the 46 genotyped individuals were subjected to isoelectric focusing and C6 phenotypes were determined after immunoblotting. Three different phenotypes were observed (Fig. 3): C6 A, C6 AB, C6 B. There was an absolute concordance between C6 protein typing and DNA typing. All seventeen individuals with genotype 22 (Glu98/Glu98) were typed as C6 A (including individual K123). All six individuals with PCR-amplified fragments being cut by DdeI only at the constant restriction site (genotype 11; Ala98/Ala98) showed protein phenotype C6 B (including individual K121). Finally, all 23 individuals with the heterozygous genotype 12 (Ala98/Glu98) were found to be C6 AB.

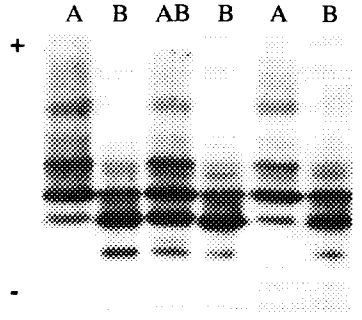

Figure 3. C6 phenotypes observed in the present study. After isoelectric focusing of plasma samples C6 protein was detected by immunoblotting. The phenotypes C6 A ('acidic') and C6 B ('basic') differ by a single step in the banding pattern; heterozygous individuals (phenotype C6 AB) show a banding pattern with two central major bands of approximately equal intensity.

Discussion

The present study demonstrates the existence of a sequence polymorphism in codon 98 of the C6 gene. Allele 1 is characterized by the nucleotide sequence GCG coding for an alanine residue, allele 2 by the nucleotide sequence GAG coding for a glutamic acid residue. This variation is easily detected as a restriction fragment length polymorphism after DdeI digestion of an amplified 112-bp portion of the C6 gene. Investigating 46 unrelated individuals absolute concordance between this RFLP and C6 protein polymorphism was observed. We thus conclude that the C6 A allotype is characterized by a glutamic acid residue and the C6 B allotype by an alanine residue at position 98 of the C6 polypeptide chain.

The point mutation in codon 98 results, at the translational level, in a charged amino acid change. With regard to isoelectric focusing, the incorporation of one negatively charged residue (glutamic acid) in C6 A should cause a shift towards the anode as compared to C6 B. Thus, this substitution might well explain the difference between C6 A and C6 B observed in isoelectric focusing gels. In fact, already in their original work on C6 polymorphism Hobart and coworkers suggested that the two common C6 allotypes differ by a single unit of charge (5). In the present study, we have demonstrated that the substitution in codon 98 consistently differentiates between C6 A and C6 B. It should also be mentioned that this substitution is the only difference between the C6 amino acid sequences reported by DiScipio & Hugli (20) and Haefliger et al. (21), respectively. Taking together, we propose that the Glu/Ala polymorphism in codon 98 is the molecular basis of the difference between the two common C6 allotypes. Nevertheless, it remains the possibility that there are other differences between *C6*A* and *C6*B* alleles. Thus, a TaqI- and an MspI-RFLP have been described using Southern analysis of genomic DNA and C6 cDNA probes (16, 27) and it will be interesting to investigate if there is linkage disequilibrium between these RFLPs and the A/B polymorphism. Investigating 46 unrelated individuals we have found an absolute concordance between codon 98 genotypes and C6 protein types. However, it seems possible that one will encounter situations showing discrepancies between DNA typing and protein typing. Such situations must not argue against our assumption that the codon 98 polymorphism is indeed the molecular basis of the two common C6 allotypes. Namely, rare C6 mutants might exist that mimic one of the common allotypes in conventional isoelectric focusing. In addition, one could imagine discrepancies due to the occurrence of *C6*Q0* alleles.

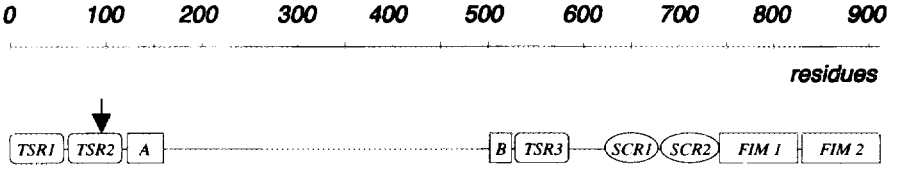

Figure 4. Structural features of human complement component C6 (modified from reference 21). The polymorphic position 98 is marked by an arrow. C6 is a typical mosaic protein; the diagram demonstrates the arrangement of the five different types of modules that have homologous counterparts in various other proteins. *TSR*: thrombospondin repeat (see text); *A*: low density lipoprotein (LDL) receptor class A module; *B*: LDL receptor class B module (epidermal growth factor precursor module); *SCR*: short consensus repeat; *FIM*: complement control factor I module. For further details see references 20, 21, 34.

Complement component C6 is a typical modular protein (20,21; Fig. 4). Little is known about structure-function relationships of C6. It has been reported that the carboxy-terminal domains of C6 (the two short consensus repeats and/or the two factor I-modules) are involved in binding to C5 (21,28,29). The polymorphic variation identified by us is located within the second thrombospondin repeat of C6. The thrombospondin module is a sequence motif, approximately 60 amino acids long, that was first identified three times in thrombospondin, an adhesive glycoprotein secreted from activated platelets (30). It is also found in other terminal complement components (C7, C8α, C8ß, C9), in properdin, a positive regulator of the alternative pathway of complement activation, and in cell surface antigens of malaria parasites (31,32, and references therein). The functional role of thrombospondin modules, in particular of those found in terminal complement proteins, is not yet characterized but could well involve interaction with membranes or protein-protein interactions (31,33). It will be interesting to explore how physicochemical properties of C6 are influenced by the charged amino acid substitution at position 98 and if this has an effect on C6 function.

Acknowledgments. We are indebted to Prof. Peter Propping for his generous support and his continued encouragement. Ulrich Weingarten and Ismet Övüc have kindly assisted with the preparation of figures. Parts of this study were supported by grants from 'Fonds der Chemischen Industrie'.

References

1. Müller-Eberhard, H.J. (1986) Annu. Rev. Immunol. 4, 503-528.
2. DiScipio, R.G. (1987) Pathol. Immunopathol. Res. 6, 343-370.
3. Morgan, B.P. (1989) Biochem. J. 264, 1-14.
4. Hobart, M.J., Lachmann, P.J., and Alper, C.A. (1975) *In* Protides of the Biological Fluids (H. Peeters, Ed.), Vol. 22, pp. 575-580, Pergamon Press, Oxford.
5. Alper, C.A., Hobart, M.J., and Lachmann, P.J. (1975) *In* Isoelectric Focusing (J.P. Arbuthnott and J.A. Beeley, Eds.), pp 306-312, Butterworths, London.
6. Tokunaga, K., Jin, F., Mauff, G., Nishimukai, H., Simeoni, E., Suzuki, K., and Weber, W. (1990) Complement Inflamm. 7, 234-239.
7. Rittner, C., Dewald, G., Berghoff, E., and Mollenhauer, E. (1979) Z. Rechtsmed. 83, 17-25.
8. Whitehouse, D.B., and Putt, W. (1983) Ann. Hum. Genet. 47, 1-8.
9. Tokunaga, K., Yamamura, N., and Omoto, K. (1984) Jpn. J. Human Genet. 29, 415-419.
10. Würzner, R., Orren, A., and Lachmann, P.J. (1992) Immunodefic. Rev. 3, 123-147.
11. Glass, D., Raum, D., Balavitch, D., Kagan, E., Rabson, A., Schur, P.H., and Alper, C.A. (1978) J. Immunol. 120, 538-541.
12. Park, K.S., Yanagisawa, M., Tokunaga, K., and Omoto, K. (1988) Jpn. J. Human Genet. 33, 325-331.
13. Hobart, M.J., and Lachmann, P.J. (1976) Transplant. Rev. 32, 26-42.
14. Hobart, M.J., Joysey, V., and Lachmann, P.J. (1978) J. Immunogenet. 5, 157-163.
15. Tokunaga, K., Dewald, G., Omoto, K., and Juji, T. (1986) Am. J. Hum. Genet. 39, 414-419.
16. Coto, E., Martinez-Naves, E., Dominguez, O., DiScipio, R.G., Urra, J.M., and Lopez-Larrea, C. (1991) Immunogenetics 33, 184-187.
17. Rogne, S., Myklebost, O., Olving, J.H., Kyrkjebo, H.T., Jonassen, R., Olaisen, B., and Gedde-Dahl Jr., T. (1991) J. Med. Genet. 28, 587-590.
18. Jeremiah, S.J., Abbott, C.M, Murad, Z., Povey, S., Thomas, H.J., Solomon, E., DiScipio, R.G., and Fey, G.H. (1990) Ann. Hum. Genet. 54, 141-147.
19. Abbott, C., West, L., Povey, S., Jeremiah, S., Murad, Z., DiScipio, R., and Fey, G. (1989) Genomics 4, 606-609.

20. DiScipio, R.G., and Hugli, T.E. (1989) J. Biol. Chem. 264, 16197-16206.
21. Haefliger, J.-A., Tschopp, J., Vial, N., and Jenne, D.E. (1989) J. Biol. Chem. 264, 18041-18051.
22. Dewald, G., Nöthen, M.M., and Cichon, S. (1993) Mol. Immunol., in press (abstract).
23. Miller, S.A., Dykes, D.D., and Polesky, H.F. (1988) Nucleic Acids Res. 16, 1215.
24. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
25. Dewald, G. (1988) Ann. Immunol. (Paris) 139, 507-515.
26. Marazziti, D., Eggertsen, G., Fey, G.H., and Stanley, K.K. (1988) Biochemistry 27, 6529-6534.
27. Coto, E., Dominguez, O., Martinez-Naves, E., Setien, F., Gutierrez, V., and Lopez-Larrea, C. (1991) Nucleic Acids Res. 19, 194.
28. Nakano, Y., Hashimoto, M., Choi, N.-H., Sugita, Y., Tobe, T., Mazda, T., and Tomita, M. (1991) Chem. Pharm. Bull. (Tokyo) 39, 432-436.
29. DiScipio, R.G. (1992) J. Biol. Chem. 267, 17087-17094.
30. Lawler, J., and Hynes, R.O. (1986) J. Cell Biol. 103, 1635-1648.
31. Goundis, D., and Reid, K.B.M. (1988) Nature 335, 82-85.
32. Nolan, K.F., Kaluz, S., Higgins, J.M.G., Goundis, D., and Reid, K.B.M. (1992) Biochem. J. 287, 291-297.
33. DiScipio, R.G., Chakravarti, D.N., Müller-Eberhard, H.J., and Fey, G.H. (1988) J. Biol. Chem. 263, 549-560.
34. Reid, K.B.M., and Day, A.J. (1989) Immunol. Today 10, 177-180.