CpG Mutations in the Reactive Site of Human C1 Inhibitor*

(Received for publication, September 1, 1988)

Karen Skriver, Elzbieta Radziejewska, Jane A. Silbermann, Virginia H. Donaldson, and Susan Clark Bock

From the Department of Microbiology/Immunology, Temple University School of Medicine and the Thrombosis Research Center, Philadelphia, Pennsylvania 19140 and the §Children's Hospital Research Foundation, Cincinnati, Ohio 45229

C1 inhibitor plays an important role in the regulation of vascular permeability through its ability to inactivate enzymes which release polypeptide kinins. Dysfunctional C1 inhibitor molecules are present in the plasma of affected members of the Da and Ri hereditary angioneurotic edema kindreds. We constructed genomic libraries from Da and Ri patient DNAs which had been cleaved with BclI to generate a fragment containing 21 kilobases of the C1 inhibitor locus. C1 inhibitor gene-containing recombinants originating from mutant Da and Ri alleles were differentiated from those derived from normal alleles by linkage analysis using the intragenic HglAI restriction fragment length polymorphism. Nucleotide sequencing of the complete protein-coding regions of the mutant alleles identified two different mutations in a CpG dinucleotide corresponding to the first two bases of the arginine codon 444. These single base mutations changed the identity of the functionally critical P1 reactive site residue from arginine to cysteine (Da) or histidine (Ri). The additional cysteine residue in C1 inhibitor Da suggests how it is covalently bound to albumin in plasma. The presence of CpG dinucleotides in the codons specifying the P1 arginines of C1 inhibitor and antithrombin III explains the high incidence of histidine and cysteine substitutions observed among dysfunctional mutants of these serine protease inhibitors.

The human plasma proteinase inhibitor C1 inhibitor plays an important role in the regulation of vascular permeability through its ability to inactivate enzymes which release polypeptide kinins. It is the most important physiological inhibitor of plasma kallikrein and factor XIa (1–5) and the only known inhibitor of C1, the first component of complement (6).

C1 inhibitor is a member of the serpin (serine protease inhibitor) gene family (7–11) which also includes the genes for other plasma protease inhibitors such as α1-antitrypsin, antithrombin III (12, 13), heparin cofactor II (14), α2-antiplasmin (15), protein C inhibitor (16), and the placent (17) and endothelial (18–20) type plasminogen activator inhibitors. Serpins are suicidal proteins which inhibit their target proteases by forming stoichiometric protease-inhibitor complexes with them. The reactive site of a serpin contains an amino acid sequence which is an ideal substrate for its target protease(s). The reactive site is located on a loop protruding from the surface of the inhibitor molecule (21, 22). In contrast to events associated with proteolysis of true substrates, serpin-protease complexes dissociate extremely slowly, and the protease remains inactive while engaged in the complex (6, 23–26). The amino acid preceding the peptide bond of the serpin that is cleaved during complex formation is called the P1 residue and is an important determinant of serpin target protease specificity. In protease-inhibitor complexes, the P1 residue is probably ester-bonded to the active site serine residue of the target protease (27).

Humans have a single gene for C1 inhibitor which maps near the centromere on chromosome 11 (8, 9). C1 inhibitor deficiency is inherited as the autosomal dominant disorder hereditary angioneurotic edema (HANE) (28). HANE patients suffer from self-limited episodes of localized, enhanced vascular permeability, particularly of subcutaneous tissue, airways, and gastrointestinal tract (29). In addition to reduced levels of serum C1 inhibitor, their C2 and C4 levels are decreased due to consumption by uninhibited, activated C2 (30, 31). Individuals with HANE are heterozygous at the C1 inhibitor locus and have one normal and one abnormal C1 inhibitor gene. Although all persons with HANE are deficient in C1 inhibitor function, the molecular defects are different in different families (32–34). Most kindreds exhibit parallel reduction of C1 inhibitor functional activity and antigen levels (type I). However, in 15–25% of HANE pedigrees, decreased functional activity is found in the presence of a dysfunctional inhibitor synthesized from the mutant allele (type II). Dysfunctional gene products often represent the predominant C1 inhibitor species in type II patient plasma due to consumption (complex formation and clearance) of the normal gene product (35–37). C1 inhibitor genes encoding three type II dysfunctional proteins did not exhibit major rearrangements compared to the normal C1 inhibitor gene (8); however, electrophoretic abnormalities and variability in the patterns of residual protease inhibitory activity of C1 inhibitor preparations from eight different type II families have been noted (38, 39). Identification of the mutations responsible for type II hereditary angioneurotic edema will improve understanding of C1 inhibitor structure/function relationships.

Dysfunctional C1 inhibitors Da and Ri have been studied previously. C1 inhibitor Da binds to albumin (32) and contains a free sulfhydryl group (40). C1 inhibitor preparations from Da patient plasma had residual protease inhibitory activities against Cls, kallikrein, activated Hageman factor, Hageman factor fragments, and plasmin which were 75, 50, 15, and 10% of normal.

The abbreviations used are: HANE, hereditary angioneurotic edema; RFLP, restriction fragment length polymorphism; kb, kilobase(s); serpin, serine protease inhibitor.
CpG Mutations in the Reactive Site of Human C1 Inhibitor

70, 194, and 20%, respectively, of the inhibitory activities displayed by an equivalent amount of normal C1 inhibitor. Preparations of C1 inhibitor from Ri patient plasma had negligible inhibitory activity against any of these proteases (38).

An HgiAI restriction fragment length polymorphism (RFLP) in the C1 inhibitor gene produces 0.7- and 0.4-kb bands on Southern blots (8). This polymorphic marker was used to identify mutant C1 inhibitor alleles in the Da and Ri pedigrees and to distinguish them from the normal C1 inhibitor alleles also present in recombinant libraries of heterozygous patient DNA. Sequencing of the mutant C1 inhibitor genes identified single base changes relative to the normal gene in both. These mutations, which have occurred in a single CpG dinucleotide encoding the first two positions of arginine 444 codon, convert the reactive site P1 residue of C1 inhibitor to a cysteine in affected members of the Da family and to a histidine in affected members of the Ri family.

MATERIALS AND METHODS2

Determination of C1 Inhibitor Functional Activity and Antigen Levels—C1 inhibitor functional activity levels in serum samples were determined by the method of Levy and Lepov (41) using the esterolytic substrate N-acetyl-L-tyrosine ethyl ester. Antigen levels were measured by rocket immunoelectrophoresis (42) or radial immunodiffusion (43).

Preparation of Genomic DNA—Peripheral blood was obtained from several members of each family, and genomic DNA was isolated as described previously (44).

Southern Blot Analysis—Three-μg samples of restriction endonuclease-treated genomic DNA were separated by agarose gel electrophoresis and transferred to nylon membranes (Gelman), which were then hybridized to 32P-nick translated C1 inhibitor cDNA probe fragments (8). In order to avoid electrophoresis and blotting artifacts, aliquots of chicken genomic DNA (which does not hydridize to the human C1 inhibitor cDNA probe) were added to samples of recombinant phage DNA used for side-by-side Southern blot analysis with normal and patient genomic DNAs.

Genomic Cloning—Family studies using the HgiAI restriction fragment length polymorphism (8) were employed to identify and mark mutant C1 inhibitor alleles in the Da and Ri families. Recombinant phage carrying the mutant alleles were obtained as follows. Genomic DNA samples (~45 μg) from HANE patients Da-II-2 and Ri-I-1, who are heterozygous for the HgiAI RFLP, were digested to completion with BclI. 18-23-kb fragments were isolated by centrifugation through 10-40% sucrose gradients, and one-fifth of the material from the appropriate fractions was inserted into the BamHI site of λDash vector DNA. These two individuals were chosen as the source of DNA for cloning experiments because they have hereditary angioneurotic edema and are heterozygotes for the HgiAI RFLP linkage marker (Fig. 1).

Three Da and six Ri recombinant phage carrying C1 inhibitor gene inserts were isolated, and their HgiAI polymorphism types were determined. Two of the Da clones carried the 0.7-kb HgiAI RFLP allele (which is associated with HANE in the Da kindred), and three of the Ri clones carried the 0.4-kb HgiAI RFLP allele (which is associated with HANE in the Ri kindred). Side-by-side Southern blot analysis of cloned phage DNA and patient and normal genomic DNAs reassured that gross rearrangement of the C1 inhibitor gene had not occurred.

RESULTS

Identification of Alleles Coding for Dysfunctional C1 Inhibitor Da and Ri Proteins—C1 inhibitor deficiency is inherited in an autosomal dominant manner, and hereditary angioneurotic edema patients are heterozygotes for a mutant (or null) C1 inhibitor gene. The HgiAI RFLP was used to follow inheritance of individual C1 inhibitor alleles in the Da and Ri HANE kindreds and to identify those carrying abnormal genes. The studies shown in Fig. 1 indicate that the dysfunctional Da gene resides on a chromosome carrying the 0.7-kb allele of the HgiAI RFLP, whereas the dysfunctional Ri gene is associated with a chromosome carrying the 0.4-kb allele.

Cloning and Sequencing—Genomic libraries were constructed from BclI digests of Da-II-2 or Ri-I-1 DNA and BamHI-digested λDash vector DNA. These two individuals were chosen as the source of DNA for cloning experiments because they have hereditary angioneurotic edema and are heterozygotes for the HgiAI RFLP linkage marker (Fig. 1).

Three Da and six Ri recombinant phage carrying C1 inhibitor gene inserts were isolated, and their HgiAI polymorphism types were determined. Two of the Da clones carried the 0.7-kb HgiAI RFLP allele (which is associated with HANE in the Da kindred), and three of the Ri clones carried the 0.4-kb HgiAI RFLP allele (which is associated with HANE in the Ri kindred). Side-by-side Southern blot analysis of cloned phage DNA and patient and normal genomic DNAs reassured that gross rearrangement of the C1 inhibitor gene had not occurred.

3 Portions of this paper (including part of “Materials and Methods,” part of “Results,” part of “Discussion,” and Figs. 4 and 5) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

Fig. 1. Inheritance of hereditary angioneurotic edema and HgiAI RFLP in Da and Ri kindreds. A, pedigrees of the Da and Ri families. Solid symbols represent HANE patients, and striped symbols represent unaffected family members. B, Southern blots prepared with HgiAI-digested genomic DNA hybridized to a 32P-labeled 500-base pair EcoRI fragment from the 3' end of human C1 inhibitor cDNA. The region of the blot containing the 0.7- and 0.4-kb HgiAI polymorphic fragments is shown. In these linkage studies, the abnormal C1 inhibitor gene segregates with the 0.7-kb HgiAI marker in the Da family and with the 0.4-kb HgiAI marker in the Ri family. Dysfunctional C1 inhibitor genes were isolated from Da-II-2 and Ri-I-1, who are heterozygous for the HgiAI marker polymorphism. C, functional activity levels of C1 inhibitor in sera determined by esterolytic assay with N-acetyl-L-tyrosine ethyl ester. Normal range for pooled sera in this assay is 6-10 units/ml. D, C1 inhibitor antigen levels in serum (mg/dl). Antigen levels were determined using the method of Laurell (42) for the Da family and by the method of Mancini et al. (43) for the Ri family. Normal range for pooled sera in these assays is 11-22 mg/dl. Patient sera used for functional and immunological assays were obtained prior to treatment with androgens.
**CpG Mutations in the Reactive Site of Human C1 Inhibitor**

**Fig. 2.** Subcloning and sequencing strategy for protein-encoding exons II–VIII of C1 inhibitor inhibitors Da and Ri. Arrowheads indicate universal (U) or C1 inhibitor gene-specific sequencing primers (see Fig. 5 for detailed description of primers). The templates for universal primer sequencing of exon V were SstI deletions of the 4.5-kb EcoRI fragment shown. E, EcoRI; S, SstI; H, HindIII; B, BamHI; N, NcoI; P, PstI. The scale bar indicates 1 kb.

**Fig. 3.** Nucleotide and derived amino acid sequences of P1 residue and surrounding reactive site region of normal human C1 inhibitor and C1 inhibitors Da and Ri. The Cpg dinucleotide in the P1 arginine codon and the single base mutations and resulting amino acid substitutions of the mutant inhibitors are indicated in **boldface letters**. The arrowhead shows the bond cleaved by C1s during complex formation.

During library construction or propagation of the DNA in bacteriophage λ.

**DISCUSSION**

An HgiAI restriction fragment length polymorphism, producing 0.7- and 0.4-kb hybridizing bands, was used to follow the inheritance of dysfunctional C1 inhibitor genes in two unrelated type II HANE kindreds, the Da and Ri families. The results from these studies are consistent with linkage of a mutant gene to the 0.7-kb allele of the HgiAI RFLP in the Da family and with linkage of a mutant gene to the 0.4-kb type HgiAI allele in the Ri family. This information was used to isolate the mutant genes from hereditary angioneurotic edema patients Da-II-2 and Ri-I-1.

In the course of this work, an efficient method for repetitive cloning of the human C1 inhibitor gene was developed. An ideal method for isolating mutant C1 inhibitor genes should satisfy several criteria. First, it should be efficient, so as to require only small amounts of patient DNA. Second, recombinants should include the entire C1 inhibitor gene locus, so that linkage analysis with the HgiAI DNA polymorphism (the only known C1 inhibitor RFLP) can be used effectively to isolate the whole mutant gene. Finally, it is also desirable that the breaking points of C1 inhibitor gene-containing inserts from different recombinants should be homogeneous in order to facilitate analysis of cloned C1 inhibitor genes from many different kindreds. These criteria suggested that an approach based on limited digestion of genomic DNA might be useful. We screened 11 restriction enzymes which produce four-base cohesive ends that are compatible with ends generated by digestion of the polylinker from commercially available λ phage vectors for the ability to produce a single, packageable-sized hybridizing band on Southern blots probed with 32P-labeled C1 inhibitor cDNA. BclI generated a single band of about 21 kb on Southern blots, and further mapping studies indicated that BclI sites are present just 5’ and 3’ to the portion of the C1 inhibitor gene which contains the protein-encoding exons (see Fig. 4). Therefore, HANE patient libraries were generated by completely digesting patient DNAs with BclI and ligating 18–23-kb fragments to BamHI arms of λDash, a bacteriophage λ cloning vector.
Universal and CT inhibitor-specific oligonucleotide primers were used to sequence the protein-encoding exons of the dysfunctional inhibitor genes. This represents a rapid method for sequencing both strands of relevant parts of the gene, and oligonucleotides developed for this project can be used readily to identify the lesions present in other mutant CT inhibitor genes as well.

Each of the dysfunctional inhibitors we studied contains a single base mutation which causes an amino acid substitution at the functionally important P1 residue. The P1 residue of normal CT inhibitor, arginine 444, has become a cysteine in CT inhibitor Da and a histidine in CT inhibitor Ri. As a consequence of these substitutions, the physiologically relevant target protease(s) of CT inhibitor may not be able to recognize or form stable, inhibitory complexes with the dysfunctional proteins, leading ultimately to a condition where regulation of proteases which release vasoactive polypeptide kinins becomes faulty, and an attack of edema occurs.

The specific P1 substitution which has occurred in CT inhibitor Ri, arginine 444 to histidine, has also been observed in a different, independently ascertained HANE family, CT inhibitor At (48). However, CT inhibitor preparations isolated from the plasma of patients from these two families displayed variable amounts of residual activity against each of five different serine proteases inactivated by CT inhibitor, with CT inhibitor Ri having negligible inhibitory activity against any of them and CT inhibitor At displaying significant residual activity (40–50% range) against three of the five proteases tested (38). Since the work reported here and the work of Aulak et al. (48) indicate that CT inhibitors Ri and At have the same P1 substitution, it is notable that their functional properties were different (38). The apparently conflicting observations may be explained by phenomena relating to contamination of some dysfunctional CT inhibitor preparations with 1) variable amounts of the patient’s normal CT inhibitor gene product and 2) trace amounts of plasma proteases, which could alter the levels of residual inhibition measured with the proteases tested from the parallel degrees of inhibition expected for copurified, uncleaved normal CT inhibitor. Alternatively, it is possible that CT inhibitor At contains another mutation in addition to the P1 histidine substitution since analysis of this mutant focused on a Pseudomonas elastase-generated C-terminal peptide.

Like CT inhibitor Ri, antithrombin III Glasgow also contains a histidine substitution at its P1 arginine residue (49). Antithrombin III Glasgow has reduced ability to inhibit thrombin, the protease target of antithrombin III, but increased affinity for heparin (50).

The mutation identified by molecular cloning and sequencing of the CT inhibitor Da gene was the substitution of P1 arginine 444 by cysteine. This result is consistent with previous work on the Da protein, in which it was shown to possess a free sulfhydryl group (40) and was isolated largely disulfide-bonded to albumin (32, 40). CT inhibitor antigen is present in the sera of Da family patients at several times the normal concentration (32). It may accumulate to abnormally high concentrations in these individuals because the rate of clearance for CT inhibitor Da complexed with albumin is very slow. Delayed clearance of another dysfunctional CT inhibitor protein (CT inhibitor Ta) has been reported by Quastel et al. (36). A phenotype similar to that of CT inhibitor Da has been described for a Swedish hereditary angioneurotic edema kindred in which those affected have CT inhibitor antigen complexed with albumin at three times the normal levels (61). The dysfunctional CT inhibitor gene from this Swedish family could have the same mutation as CT inhibitor Da.

The P1 arginine to cysteine mutation observed in CT inhibitor Da is also found in the Northwick Park variant of the homologous serpin antithrombin III (49). As is the case for CT inhibitor Da, antithrombin III Northwick Park is also complexed with albumin (52); however, antithrombin III antigen in Northwick Park patients does not accumulate to the very high levels observed for CT inhibitor antigen in Da patients (53).

A third P1 arginine to cysteine serpin mutant has been studied. The gene for α1-antitrypsin cysteine 358 was generated in vitro mutagenesis techniques, and the mutant protein was expressed in yeast (54). Mutant protein which had been preincubated with 1 mM dithiothreitol reacted rapidly with pancreatic elastase and neutrophil elastase (the physiological target of α1-antitrypsin), and the second-order association rates were similar to those of the native inhibitor containing a P1 methionine. Purified mutant protein which was not pretreated with reducing agent migrated as a dimer on sodium dodecyl sulfate-polyacrylamide gel electrophoresis and was not an effective inhibitor. These findings suggest that angioedema in the Da family and thrombosis in the Northwick Park family may result from complex formation between the respective serpins and albumin, rather than from an inherent lack of inhibitory activity in these serpin mutants.

Alternatively, the introduction of a fifth cysteine in CT inhibitor Da and a seventh cysteine in antithrombin III Northwick Park may lead to incorrect disulfide bond formation and aberrant folding of these serpin variants.

A large number of P1 mutations are to be expected among serpin dysfunctional mutants. However, it is notable that only two of the six possible amino acid substitutions which could theoretically occur at the P1 arginines of CT inhibitor and antithrombin III have actually been observed. The apparently high incidence of P1 mutations and of recurring histidine and cysteine substitutions in particular is consistent with the presence of a CpG dinucleotide in the first two positions of the codons specifying arginine 444 (CGC) in CT inhibitor and arginine 393 (CGT) in antithrombin III. 5-Methylcytosine-containing CpG dinucleotides are the major methylated sequences in vertebrates, and a large body of evidence supports the hypothesis that they are hotspots for mutation via a mechanism involving spontaneous deamination of 5-methylcytosine to thymine (55–57). For instance, CpG dinucleotides are present much less frequently than expected in sequenced human genes (observed/expected = 0.37) (58). Also, restriction endonuclease recognition sites containing CpG dinucleotides are known to exhibit increased rates of polymorphism in human DNA (59). Finally, 35% of single base mutations causing human genetic disease occur within CpG dinucleotides (60).

Thus, we believe that the molecular mechanism underlying the P1 arginine to cysteine mutations in CT inhibitor Da and antithrombin III Northwick Park is deamination of the C in the first position of CGY arginine codons to generate TGY cysteine codons (where Y = pyrimidine). Similarly, deamination of noncoding strand C residues in the second position of the arginine-encoding CGY codons of CT inhibitors Ri and At and antithrombin III Glasgow would result in the substitution of A residues (and histidine-encoding CAY codons) on the sense strand. The P1 residues of six human serpins are arginines. In CT inhibitor, antithrombin III, α1-antitrypsin, and endothelial plasminogen activator inhibitor, the P1 arginines are encoded by CGY codons (15, 18–20, 61). The P1 arginines of human protein C inhibitor (16) and placental plasminogen activator inhibitor (17) are AGR (where R = purine). Or: the basis of the information discussed above, one
would predict that a high incidence of cysteine and histidine substitutions will be found at the PI arginines of C inhibitor, antithrombin III, α2-antiplasmin, and endothelial plasminogen activator inhibitor and that the incidence of PI substitutions observed in naturally occurring protein C inhibitor and plasminogen activator inhibitor mutants will be much lower due to the absence of CpG in the AGR codons specifying their reactive site arginines.

Acknowledgments—We thank members of the Da and Ri families for their participation in this study and Dr. P. Model for helping us to synthesize oligonucleotides.

REFERENCES
CpG Mutations in the Reactive Site of Human C1 Inhibitor

Understanding the structure and regulation of the normal C1 inhibitor gene will facilitate investigation of variant C1 inhibitor genes. A single base change in the human C1 inhibitor gene and two in the transcribed region of tissue pool at the non-transcribed portion of each promoter.

Materials and Methods

Adequate levels of recombinant human recombinant results in C1 inhibitor gene were obtained with 5' and 3' leader sequences generating antisense probes. Oligonucleotide solutions (10 and 100 nM) were used to generate antisense and be cloned into the 5' leader. The results obtained with the normal C1 inhibitor gene were obtained with 5' and 3' leader sequences generating antisense probes (Fig. 1). The antisense probes were obtained with the normal C1 inhibitor gene were obtained with 5' and 3' leader sequences generating antisense probes (Fig. 1).

Results and Discussion

Five oligonucleotide probes overlapping fragments of the human C1 inhibitor gene were obtained by sequencing DNA fragments of the library described above. Southern blot analysis of the cloned BamHI and genomic DNA samples hybridized in adjacent lanes of the same agarose gel revealed that some internal fragments of two recombinants did not correlate with the corresponding genomic bands. We believe these discrepancies were the result of its potential for hybridizing to sequences from the promoter region of human recombinant human C1 inhibitor gene.

The top panel of the C1 inhibitor gene presented in Fig. 1 shows the enzyme activity in human. The gene is transcribed from the promoter region of human recombinant human C1 inhibitor gene.

Figure 1

A schematic map of the human C1 inhibitor gene. The 5' leader (BamHI E) and 3' leader (BamHI F) are shown. Exons are numbered as in Carter et al. (1988) and are shown. The map contains the promoter region of human recombinant human C1 inhibitor gene.