Patients with Goodpasture’s disease have two normal COL4A3 alleles encoding the NC1 domain of the type IV collagen α3 chain

Ulf Persson1, Jens Michael Hertz2, Malin Carlsson1, Thomas Hellmark1, Inger Juncker2, Jörgen Wieslander3 and Märten Segelmark1

1Department of Nephrology, Lund University, 3Wieslab AB, Lund, Sweden and 2Department of Clinical Genetics, Aarhus University Hospital, Aarhus, Denmark

Abstract

Background. Goodpasture’s disease (GP) is a rare but severe disease characterized by anti-glomerular basement membrane antibodies, rapidly progressive glomerulonephritis and lung haemorrhage. The autoantibodies are restricted to a narrow epitope region on the NC1 domain of the α3 chain of type IV collagen. GP is strongly associated with major histocompatibility complex (MHC) allele HLA DRB1-15. Recent research, however, has failed to identify a T-cell epitope with molecular characteristics that explain the relationship between the MHC class II molecule and the autoantibody generation. We hypothesized that an as yet unidentified sequence variant in exons 48–52 of the COL4A3 gene that encodes the NC1 domain of the type IV collagen α3 chain could generate a new peptide sequence that, through interaction with specific MHC class II molecules, would increase the risk of developing GP.

Methods. All patients previously treated for GP at the Lund and Malmö University Hospitals, who were alive at the time of the study, were asked to participate. DNA was extracted from leukocytes and subjected to genomic tissue typing and sequencing of the COL4A3 gene exons 48–52.

Results. All 15 patients in the study had a nucleotide sequence in the COL4A3 gene encoding a protein identical to GenBank entry NM_000091. HLA D allele distribution was in line with previous publications, showing a strong positive association between HLA DRB1-15, HLA DQB1-6 and GP (P < 0.02). Of the 15 GP patients, 73% carried HLA DRB1-15 and 87% carried the HLA DQB1-6 antigen. Corresponding figures for the controls were 27 and 50%.

Conclusion. This study effectively falsifies the hypothesis that a minor alteration in the COL4A3 gene could be a major factor in the aetiology of GP. Scandinavian GP patients have an MHC distribution similar to that which has been described previously for Anglo-Saxon patients.

Keywords: COL4A3 gene/Goodpasture’s disease

Introduction

Goodpasture’s disease (GP) is characterized by the combination of rapidly progressive glomerulonephritis and anti-glomerular basement membrane (anti-GBM) antibodies, and is often accompanied by lung haemorrhage [1]. It usually results in permanent renal failure or death if not diagnosed and treated promptly. Transfer experiments have proven the toxicity of the antibodies in primates, and the removal of the antibodies is part of successful treatment of patients [2,3]. The antibodies are directed to the C-terminal non-collagenous domain (NC1) of type IV collagen. Most antibodies, maybe all disease-causing antibodies, react with a limited number of epitopes on the NC1 domain of the α3 chain [4,5]. Of special interest is an epitope region in the N-terminal third of the domain, where amino acid residues between the positions T1455 and S1469 together with Q1495 seem to be critical for the binding of the autoantibodies [6,7]. Antibodies reacting with other epitope regions and antibodies cross-reacting with other α-chains can also be found in many patients, but their pathogenic and diagnostic potentials remain unproven. Recent work regarding the three-dimensional structure of the NC1 domain and the quaternary organization of the α-chains has shown that GBM NC1 domains consist of two head-to-head-oriented α3,4,5 heterotrimers [8]. The
patients with Goodpasture’s disease have two normal COL4A3 alleles.

While the B-cell epitopes are well characterized, less is known about T-cell involvement. The IgG subclass distribution of the autoantibodies, with a preponderance for subclasses IgG1 and IgG4, suggests an antigen-driven and T-cell-dependent process [9]. Studies of the functional affinity, however, have been unsuccessful in establishing evidence for affinity maturation during the course of the disease [10]. The strongest indirect evidence for T-cell involvement in the pathogenesis stems from the skewed distribution of major histocompatibility complex (MHC) class II alleles among patients with GP, compared with controls. Several studies have shown strong positive association with the DR alleles HLA DRB1-15 and DRB1-4, and negative associations with HLA DRB1-7 and DRB1-1 [11]. HLA DRB1-15 alleles have been found in 70–90% of patients of Western European descent, compared with 20–30% in the background population. Phelps and co-workers have launched extensive studies in order to elucidate specific T-cell epitopes of importance [12–14]. They found that none of the overlapping peptides they constructed bound HLA DRB1-15 with much higher affinity than for both DRB1-1 and DRB1-7. This included peptides previously identified as being expressed by HLA DRB1-15-positive antigen-presenting cells (APCs) after processing of recombinant z3(IV) [13]. Instead, the tendency was in the other direction, and they proposed that DRB1-1/7 might protect from anti-GBM disease by capturing peptides in heterozygotic APCs. In a recent article from the same group, it was shown that T-cells from patients, in contrast to controls, showed reactivity towards certain z3(IV)-derived peptides presented by HLA DRB1-15 [14]. However, the relationship between HLA DRB1-15 and GP still remains unexplained.

It is well established that inherited defects in the COL4A3 gene can cause renal disease. A homozygotic state of non-functioning alleles or alleles coding for non-functioning proteins gives rise to the autosomal recessive form of Alport syndrome, while a heterozygotic state usually manifests as thin basement membrane disease [15]. After renal transplantation, some patients with Alport syndrome develop anti-GBM antibodies, which can lead to transplant failure [16]. Thus, defects in type IV collagen genes may under certain circumstances predispose to anti-GBM production.

While HLA DR1-15 is common and GP is rare, other factors, genetic or environmental, must be critical for the development of anti-GBM antibodies. We hypothesized that there might exist sequence variants in the COL4A3 gene leading to alterations in the primary amino acid sequence of the NC1 domain. This could subsequently lead to the generation of new T-cell epitopes having the ability to interact unfavourably with specific MHC alleles, increasing the risk of developing anti-GBM antibodies. In order to test this hypothesis, we sequenced exons 48–52 of the COL4A3 gene in 15 patients who previously had been treated for GP and determined their MHC class II genotype. The findings did not support the hypothesis.

**Patients and methods**

Databases of biopsy registries at the University Hospitals of Lund and Malmö, and the Wieslab AB serological database were searched for patients with GP and anti-GBM antibodies. Patients were included if they had circulating anti-GBM antibodies that could be confirmed by retesting stored sera, and had a clinical history indicating rapidly progressive glomerulonephritis such as haematuria, casts in the urinary sediment and a rising serum creatinine concentration. All patients alive at the time of the study were contacted and asked to participate. After informed consent was granted, clinical data were retrieved from the patients’ hospital records. The local ethics committee approved the protocol.

**HLA typing**

Leukocytes were prepared from whole blood by density centrifugation, and DNA was extracted as described by Miller et al. [17]. HLA DRB1, DRB3-5 and DQB1 typing was done at the local tissue-typing laboratory using polymerase chain reaction (PCR)-based standard clinical procedures. Results were compared with previously established allele frequencies based on the local population of blood donors.

**PCR amplification of genomic DNA**

Primers for COL4A3 exons 48–52 were synthesized according to published sequences [GenBank accession no. NM_000091 (http://www.ncbi.nlm.nih.gov/)] (Table 1). Genomic DNA was amplified in a reaction volume of 25 μl containing 200 ng of DNA, 200 μM dNTP, 20 pmol of each primer and 1 U of Taq DNA polymerase (Boehringer Mannheim). The buffer used for PCR contained 10 mM Tris–HCl, pH 8.3, 50 mM KCl and 1.5 mM MgCl2. The PCR conditions were as follows: initial denaturation at 94°C for 4 min followed by 30 cycles of denaturation at 94°C for 1 min, annealing for 1 min at the temperatures as indicated in Table 1, and extension at 72°C for 2 min. The final extension was 4 min at 72°C.

**Sequencing of PCR products**

Sequencing was performed by cycle sequencing of double-stranded DNA using an ABI PRISM Big Dye Terminator Labelling Cycle Sequencing Kit (PE Applied Biosystems) as recommended by the manufacturer, and run on an ABI PRISM 310 Genetic Analyzer (PE Applied Biosystems). Obtained sequences were analysed by Sequence Navigator (Applied Biosystems) and compared with control sequences from a normal individual. Nucleotides are numbered starting from the A of the ATG initiator codon.

**ELISA**

An enzyme-linked immunosorbent assay for circulating antibodies against z3(IV) NC1 (anti-GBM), proteinase 3
[PR3-anti-neutrophil cytoplasmic antibody (ANCA)] and myeloperoxidase (MPO-ANCA) was carried out as described previously [9,18]. In brief, antigen was coated in microtitre plates at a concentration of 0.5 \( \mu \)g/ml for \( \alpha(3) \) NC1, and at 1 \( \mu \)g/ml for PR3 and MPO in carbonate buffer. After washing, plates were incubated for 1 h with sera diluted 1:100 in ELISA buffer. In the next step, alkaline phosphatase-conjugated goat anti-human IgG (Sigma, St Louis, MO) was incubated on the plates at a dilution of 1:20 000. Finally, a substrate buffer (1 M diethanolamine, 0.5 mM MgCl\(_2\), pH 9.8) with \( p \)-nitrophenyl phosphate was added and the optical density (OD) at 405 nm was read with a spectrophotometer. All samples were run in triplicate and the results were expressed as arithmetic means. A positive result was defined as being above the mean results of healthy controls +3SD (0.19 for GP, 0.16 for PR3 and 0.13 for MPO).

### Results

For this study, 15 patients were identified who had been treated previously for anti-GBM-mediated renal disease in Lund or Malmö. All patients agreed to participate in the study. Of the 15 patients, seven were men and eight were women. The mean age at diagnosis was 42.8 years (range 18–68, Table 2). All patients had circulating anti-GBM antibodies at diagnosis, and in 11 of the patients kidney-reactive antibodies were also detected by direct immunofluorescence. There was a wide range in kidney function at the time of diagnosis as reflected by serum creatinine concentrations ranging from 181 to 1240 \( \mu \)mol/l (median value 608).

Sequencing of exons 48–52 of the COL4A3 gene yielded a sequence identical to the published sequence in 14 of the 15 patients. In patient no. 9, a single base substitution at position 5301 was identified (5301 C>T). This is 288 bp 3' of the TGA stop codon in exon 52 and does not result in any change in the amino acid composition.

Tissue typing of the MHC locus HLA D revealed a significantly skewed pattern (Tables 3 and 4). Eleven of the 15 patients (73%) were positive for DRB1-15 compared with 27% in the background population (\( P<0.001 \)). All 11 DRB1-15-positive patients were also positive for DQB1-06. In total, 13 out 15 patients (87%) were positive for DQB1-06 as compared with 50% among the blood donors (\( P<0.01 \)).

### Table 1. Primer sequences and conditions for PCR amplification of the 3' end of COL4A3 encoding the NC1 domain of the \( \alpha(3) \) chain of type IV collagen

<table>
<thead>
<tr>
<th>Exon</th>
<th>Forward primer (5'–3')</th>
<th>Reverse primer (5’–3’)</th>
<th>Annealing temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>48</td>
<td>GCTGGATCATGGAGTGGCC</td>
<td>TGCATGTGGAACACGACAA</td>
<td>55</td>
</tr>
<tr>
<td>49</td>
<td>GCTTTGTGTGTAGTAACGATGC</td>
<td>CCATTACAAAGGAACTGATTC</td>
<td>55</td>
</tr>
<tr>
<td>50</td>
<td>TGCGTCAACACACCTCTCCTC</td>
<td>TGTACATTCTCAGGATGTC</td>
<td>55</td>
</tr>
<tr>
<td>51</td>
<td>TGAACCCCAATGGGACAGAGT</td>
<td>TGAATAGTTCGTCAATTGATGC</td>
<td>50</td>
</tr>
<tr>
<td>52</td>
<td>CAGCAGAAATTTCCCTTCTATGC</td>
<td>AATGCTCCAGGATACACAT</td>
<td>55</td>
</tr>
</tbody>
</table>

Primer sequences are synthesized according to published sequence data (GeneBank accession no. NM_000091) (http://www.ncbi.nlm.nih.gov/). Exonic sequences are underlined. The number of intron bases between primer and exon are indicated in parentheses.

### Table 2. Patient characteristics

<table>
<thead>
<tr>
<th>Patient number</th>
<th>Sex</th>
<th>Age at diagnosis (years)</th>
<th>Biopsy findings</th>
<th>Creatinine at diagnosis (( \mu )mol/l)</th>
<th>Follow-up (years)</th>
<th>Creatinine at follow-up (( \mu )mol/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>17</td>
<td>ND</td>
<td>333</td>
<td>15</td>
<td>114</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>19</td>
<td>Cres, lin</td>
<td>455</td>
<td>16</td>
<td>Tp</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>19</td>
<td>Cres, lin</td>
<td>335</td>
<td>17</td>
<td>Tp</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>19</td>
<td>Cres, ND</td>
<td>835</td>
<td>14</td>
<td>Tp</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>24</td>
<td>Cres, lin</td>
<td>1240</td>
<td>13</td>
<td>Tp</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>28</td>
<td>Cres, lin</td>
<td>1310</td>
<td>16</td>
<td>Tp</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>39</td>
<td>Cres, lin</td>
<td>206</td>
<td>5</td>
<td>109</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>41</td>
<td>Cres, lin</td>
<td>1200</td>
<td>9</td>
<td>HD</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>51</td>
<td>Cres, ND</td>
<td>608</td>
<td>4</td>
<td>HD</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>53</td>
<td>Cres, lin</td>
<td>824</td>
<td>8</td>
<td>Tp</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>56</td>
<td>Cres, lin</td>
<td>584</td>
<td>7</td>
<td>Tp</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>64</td>
<td>Cres, lin</td>
<td>975</td>
<td>0.5</td>
<td>HD</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>65</td>
<td>Cres, unspecific</td>
<td>181</td>
<td>10</td>
<td>168</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>66</td>
<td>Cres, lin</td>
<td>475</td>
<td>1</td>
<td>HD</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>68</td>
<td>Cres, lin</td>
<td>626</td>
<td>5</td>
<td>160</td>
</tr>
</tbody>
</table>

Cres = crescentic glomerulonephritis; lin = linear immunofluorescence for IgG along the GBM; ND = not done; HD = haemodialysis; Tp = renal transplant.
DQB1-06-negative patients were both positive for DRB1-04 and DQB1-03. No patients were found with any of the following alleles: DRB1-07, 08, 09, 10, 11 or 12. Individually, none of these negative associations is statistically significant.

Five patients were ANCA positive in ELISA, four in MPO-ANCA ELISA and one in PR3-ANCA ELISA. ANCA-positive patients were older than ANCA-negative patients, median age 65 vs 26 years. This was highly significant; the Mann–Whitney U-test yielded a P-value of 0.0013. There were no trends suggesting a different MHC genotype among ANCA-positive patients. Four of the five ANCA-positive patients were DRB1-15 positive, and both DQB1-06-negative patients were ANCA negative. MHC genes did not seem to influence the anti-GBM titre. The mean OD value for the 11 DRB1-15-positive patients was similar to the OD value for the four DRB1-15-negative patients (1.8 vs 1.6, P = 0.47).

### Discussion

This study effectively falsifies the hypothesis that minor alterations in the Goodpasture antigen are a major factor in the aetiology of GP. We reasoned that if a sequence variant would cause a new T-cell epitope, forming a specific interaction with a particular MHC antigen, this variant had to be present in > 30% of the DRB1-15 cases in order to be considered a major factor. We found no alteration in the gene encoding the primary amino acid sequence among 11 DRB1-15 carriers. The probability that such an association still exists in spite of our findings is very low (P < 0.02). A second possibility is that a certain variant may make DRB1-15 non-carriers susceptible to anti-GBM production. Even though we found no variants among the four non-carriers in the present study, we cannot exclude the possibility that such variants exist in a majority of the non-DRB1-15 cases, but it is unlikely...
It has been proposed that the GP epitope is not normally exposed to the immune system and that environmental factors such as hydrocarbon exposure may unmask the epitope [22]. It has also been suggested that reactive oxygen species can alter the structural integrity of the GBM and thereby increase its auto-immunicity [23]. Even though these theories might have bearing on the aetiology of GP, they all fail to explain the strong MHC association. Recent results from Wong and co-workers showing that the α3(IV) and α5(IV) chains are expressed in the thymus (without incorporation into a basement membrane) [24] open up another line of thinking. This can be interpreted as if GP epitope-reactive T cells normally are kept under control by central tolerance, and that the MHC association might stem from a defective clonal deletion due to the low affinity between DRB1-15 and NC1 peptides [12]. The production of NC1 by non-basement membrane-producing cells also opens up the possibility that other cells, including those of a haematopoietic origin, under pathological conditions might transcribe the COLA43 gene, eventually leading to autoantibody production.

In conclusion, we found that GP among Scandinavian patients is strongly linked to HLA DRB1-15, and that this association is not a consequence of interaction with certain genetic variants of the GP antigen.

Acknowledgements. This work was supported by grants from the Swedish Scientific Research Council and by the Swedish Kidney Foundation.

Conflict of interest statement. None declared.

References

8. Borza DB, Bondar O, Todd P et al. Quaternary organization of the Goodpasture autoantigen, the alpha 3(IV) collagen chain. Sequestration of two cryptic autoepitopes by intrapromoter
Patients with Goodpasture’s disease have two normal COL4A3 alleles.

Interactions with the alpha4 and alpha5 NC1 domains. J Biol Chem 2002; 277: 40075–40083


13. Phelps RG, Jones VL, Coughlan M, Turner AN, Rees AJ. Presentation of the Goodpasture autoantigen to CD4 T cells is influenced more by processing constraints than by HLA class II peptide binding preferences. J Biol Chem 1998; 273: 11440–11447


Received for publication: 28.11.03
Accepted in revised form: 31.3.04