Mutations in the PIG-A Gene Causing Paroxysmal Nocturnal Hemoglobinuria Are Mainly of the Frameshift Type

By Khedoudja Nafa, Philip J. Mason, Peter Hillmen, Lucio Luzzatto, and Monica Bessler

Paroxysmal nocturnal hemoglobinuria (PNH) is an acquired hemolytic anemia associated with somatic mutations in the X-linked gene PIG-A, which encodes a protein involved in the biosynthesis of glycosylphosphatidylinositol anchors. To further elucidate the molecular basis of paroxysmal nocturnal hemoglobinuria, we have worked out a systematic and relatively rapid methodology to scan for mutations in the entire coding region of the PIG-A gene. By this methodology, we have identified 15 different somatic mutations in 12 patients. The mutations were spread throughout the entire PIG-A-coding region. Of the mutations, 10 caused frameshifts, 6 caused small deletions, 3 caused small insertions, and 1 caused deletion-insertion. Five single base pair substitutions caused three missense mutations, one nonsense mutation, and one defect in the donor splice site of intron 4. In each of 3 patients, two independent mutations were identified. The predominance of frameshift mutations may reflect selection for somatic mutations giving rise to clones with a completely nonfunctional PIG-A protein.

© 1995 by The American Society of Hematology.

Materials and Methods

Patients. A total of 17 PNH patients (6 males, 11 females) were investigated. The diagnosis of PNH was based on the presence of hemolytic anemia and a positive Ham test. In 6 patients, severe aplastic anemia, according to the definition of the Working Party on Severe Aplastic Anemia of the European Group of Bone Marrow Transplantation, with an initially negative Ham test was diagnosed before PNH (AA-PNH).

Quantification of GPI-linked surface proteins. Analysis of GPI-anchored membrane proteins on RBCs, granulocytes (polymorphonuclear cells [PMN]), and lymphocytes was performed as previously described by flow cytometry (FACscan; Becton Dickinson, Mountain View, CA) using monoclonal antibodies towards GPI-linked surface antigens (CD59, CD55, decay-accelerating factor [DAF], CD16, and CD48). Cells in which GPI-linked proteins are completely absent were designated as PNH III; cells in which GPI-linked proteins were significantly decreased were designated as PNH II.

PNH cell lines. Production and analysis of Epstein-Barr virus-transformed lymphoblastoid cell lines has been described elsewhere.

DNA extraction. Granulocytes and mononuclear cells were separated by centrifugation over a Ficoll gradient. RBCs were removed by hypotonic lysis. Genomic DNA was extracted separately from peripheral blood granulocytes and mononuclear cells by the sodium dodecyl sulfate proteinase-k method. When necessary, blood samples were shipped from other centers in adenine citrate dextrose (ACD) at room temperature and were processed within 48 hours. Alternatively, we obtained Ficoll-separated frozen cells or DNA extracted individually from granulocytes and mononuclear cells.

Characterization of PIG-A mutations. We undertook a systematic approach whereby the coding sequence of the PIG-A gene was amplified from genomic DNA in four different portions using intronic primers (Fig 1). Polymerase chain reaction (PCR) amplification, single-strand conformation analysis (SSCA), and heteroduplex analysis (HA) were performed as previously described.

Nucleotide (nt) sequencing. Nt sequencing was performed using the Sequenase Version 2.0 sequencing kit (US Biochemical Corp, Cleveland, OH). Nt sequencing was performed either by direct sequencing (DS) using biotinylated primers or by cloning into plage M13. For the latter, 7 to 24 M13 clones of each amplification product were sequenced each time. To facilitate detection of nt changes, sequencing reactions were loaded track-by-track (see Fig 2). To exclude PCR-artefacts, only changes in the PIG-A nt sequence found in at least 2 individual M13 clones were counted. If the mutation caused a change in a restriction site, the mutation was verified by the appropriate restriction enzyme digest. DNA from the patient’s mononuclear cells served as an internal control to show that the mutation was not constitutional.

Results

Phenotype analysis. Of the 17 patients we have investigated (6 males and 11 females), 11 had primary, or classic, PNH, whereas in 6 (2 males and 4 females) the diagnosis of AA had been made before that of PNH (AA-PNH). With respect to RBCs, flow cytometry analysis showed that 2 of 17 patients had only PNH II cells. In 7 patients, both PNH II and PNH III cells were observed, whereas in 8 patients...
only PNH III cells were present. The percentage of granulocytes with the PNH phenotype varied between 20% and 100% (see Table 1).

DNA analysis. To localize the mutation within the 1452-bp coding region of the PIG-A gene, we used HA and SSCA of PCR-amplified genomic DNA. Four and three mutations were localized to a specific DNA fragment by SSCA and by HA, respectively, whereas 8 were seen by both methods (Fig 3). One mutation (HH15) was found through the loss of a recognition site for a restriction enzyme that we used for both SSCA and HA. In these 12 patients, the shifted DNA fragments were further analyzed by nt sequencing, leading to the identification of a total of 15 mutations (Table 1). We did not succeed in localizing the mutation in 5 patients by following this algorithm.

A total of 4 patients had small deletions, and 2 had small insertions. In all these cases, the effect of the mutations was to cause a frameshift in the PIG-A coding sequence, leading to a truncated protein. Three patients were found to have single base pair changes (2 single amino acid [aa] substitutions; 1 early stop codon; see Table 1). In each of 3 patients we found two mutations. In HH50, the two mutations, a single base-pair substitution and a deletion/insertion causing a frameshift, were remarkably close within the gene. Because they were found on independent M13 clones, we can be certain that they were in different cell populations (Fig 4). The same applies to the two mutations in patients HH24 and HH54.

To confirm that the mutations observed had arisen as somatic mutations, we tested for their presence in DNA from mononuclear cells, in which the PNH component is usually smaller than that of granulocytes and, therefore, can serve as individual internal controls. This criterion was met in all of the 15 mutations.

Constitutional changes in the PIG-A gene. In HH13 (female PNH patient) we found, in addition to the somatic mutation, a second change in the PIG-A gene. In contrast to the former, the latter was equally present in the patient’s granulocytes and mononuclear cells. Sequence analysis showed that HH13 was heterozygous for a cytidine-to-thymidine substitution at position nt 55, causing a single aa change from arginine to tryptophan at position aa 19. This single
Fig 2. Nt sequence of M13 clones containing exon 5 of patient HH54. Sequencing reactions were loaded track-by-track (G, A, T, and C) to facilitate the detection of nt changes. The insertion of T in 3 of 8 M13 clones is easily detected.

base pair substitution causes the loss of the Ava I restriction site at position nt 53. Two additional PNH patients (HH22 and HH55) were also found to be heterozygous for the loss of the Ava I restriction site. Although the aa change from arginine to tryptophan seems rather drastic, the PIG-A pro-

**Table 1. Clinical Manifestations, PNH RBC Phenotype, and Molecular Characterization of PIG-A Mutations in 12 Patients With PNH**

<table>
<thead>
<tr>
<th>UPN</th>
<th>Sex</th>
<th>Clinical Type</th>
<th>RBC PNH Type</th>
<th>PMN PNH Cells (%)</th>
<th>Exon Method</th>
<th>bp Change</th>
<th>aa Change and Size of Mutative PIG-A Protein (aa)</th>
<th>M13 Clones (Mutant/Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH2</td>
<td>M</td>
<td>Classic</td>
<td>III</td>
<td>85</td>
<td>3</td>
<td>SSCA</td>
<td>Deletion 820 G Frameshift (289)</td>
<td>DS</td>
</tr>
<tr>
<td>HH4</td>
<td>M</td>
<td>Classic</td>
<td>III</td>
<td>100</td>
<td>3</td>
<td>HA + SSCA</td>
<td>Insertion 761 A Frameshift (281)</td>
<td>DS</td>
</tr>
<tr>
<td>HH12</td>
<td>F</td>
<td>AA-PNH</td>
<td>II</td>
<td>100</td>
<td>2</td>
<td>SSCA</td>
<td>548 TGT → TTT 183 cysteine → phenylalanine</td>
<td>3/7</td>
</tr>
<tr>
<td>HH13*</td>
<td>F</td>
<td>Classic</td>
<td>III + II</td>
<td>93</td>
<td>2</td>
<td>HA</td>
<td>142 GCC → AGC 48 glycine → serine</td>
<td>DS</td>
</tr>
<tr>
<td>HH15*</td>
<td>F</td>
<td>Classic</td>
<td>III + II</td>
<td>83</td>
<td>5</td>
<td>HA + SSCA</td>
<td>55 CGG → TGGT 19 arginine → threonine</td>
<td>DS</td>
</tr>
<tr>
<td>HH24</td>
<td>M</td>
<td>Classic</td>
<td>III + II</td>
<td>95</td>
<td>2</td>
<td>HA + SSCA</td>
<td>Deletion 1132 C Frameshift (379)</td>
<td>DS</td>
</tr>
<tr>
<td>HH27</td>
<td>M</td>
<td>Classic</td>
<td>III</td>
<td>97</td>
<td>3</td>
<td>HA + SSCA</td>
<td>Deletion 770 T Frameshift (256)</td>
<td>5/7</td>
</tr>
<tr>
<td>HH33</td>
<td>F</td>
<td>Classic</td>
<td>III + II</td>
<td>95</td>
<td>2</td>
<td>HA + SSCA</td>
<td>Deletion 118 G Frameshift (69)</td>
<td>8/22</td>
</tr>
<tr>
<td>HH40</td>
<td>F</td>
<td>AA-PNH</td>
<td>III</td>
<td>83</td>
<td>2</td>
<td>SSCA</td>
<td>298 CAG → TAG 100glutamine → AMB</td>
<td>4/12</td>
</tr>
<tr>
<td>HH60</td>
<td>F</td>
<td>AA-PNH</td>
<td>III</td>
<td>96</td>
<td>2</td>
<td>HA + SSCA</td>
<td>384 (TAGTTC) → A Frameshift (134)</td>
<td>7/24</td>
</tr>
<tr>
<td>HH53</td>
<td>F</td>
<td>Classic</td>
<td>III + II</td>
<td>97</td>
<td>2</td>
<td>HA</td>
<td>387 AGT → AGA 129 serine → arginine</td>
<td>4/24</td>
</tr>
<tr>
<td>HH64</td>
<td>M</td>
<td>AA-PNH</td>
<td>III</td>
<td>85</td>
<td>4 + 5</td>
<td>HA</td>
<td>Insertion 520 GT Frameshift (194)</td>
<td>4/11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Insertion 1060 A Frameshift (356)</td>
<td>3/11#</td>
</tr>
</tbody>
</table>

* In this case, the analysis was performed on a PNH III-lymphoblastoid cell line (HH13/49P). The somatic mutation found was confirmed by sequencing genomic DNA from the patient’s granulocytes.

† HH13 is heterozygous for the inherited 55 C-to-T substitution.

‡ In this case, we analyzed PNH III-lymphoblastoid cell line (HH15/3P) because no granulocyte DNA was available.16

§ The 8 M13 clones with the deletion 189 A did not have the deletion 680 C and vice versa.

‖ The 7 M13 clones with deletion-insertion did not have the T-to-A substitution found in the other 4 M13 clones and vice versa.

¶ The same 5' splice site mutation (CAGgt → CAGgt) was found to affect RNA splicing in a Japanese patient.8 Splicing of exon 3 to exon 5 leads to a truncated open reading frame of 285 codons, which is a sufficient explanation for the lack of GPI-linked proteins on the surface of RBCs and granulocytes.

# The g-to-a substitution and the insertion 1060 A were on different M13 clones.

Fig 3. SSCA and HA of the PIG-A coding region. (A) SSCA of exon 2 digested with Taq I and HindIII from patient HH13. HH13 is a female PNH patient in whom analysis was performed from genomic DNA of an Epstein-Barr virus-transformed lymphoblastoid cell line deficient in GPI-anchors (HH13/49P). In comparison with the 365-bp Taq I fragment from the phenotypically normal cell line (HH13/30N), we find in HH13/49P an altered electrophoretic mobility (shift) of about half of her corresponding DNA fragment. (B) HA of exon 3 in patient HH4, in whom the entire granulocyte population had the PNH phenotype. Only mutant homoduplex molecule was formed, detected by altered electrophoretic mobility as compared with that of a normal control. (C) HA of exon 3 in patient HH27. The multiple fragments are caused by the different electrophoretic mobility of normal homoduplex, mutant homoduplex, and heteroduplex.
tein with tryptophan at position aa 19 does not cause the PNH phenotype. Indeed, we observed the constitutional loss of the Ava I site in 3 males not suffering from PNH, in one of whom we documented a negative Ham test and a normal expression of GPI-linked proteins on blood cells. The constitutional loss of Ava I has recently been reported independently by Endo et al.18

DISCUSSION

Screening for PIG-A mutations. Our aim was to develop a generally applicable strategy, which, in contrast to the previously used methodologies,3,7 would also be suitable for identifying unknown PIG-A mutations in stored patient samples and in samples sent from long distances. The method described here is based on SSCA and HA of PCR-amplified genomic DNA fragments. We were mindful of the fact that (1), because the PIG-A gene is located on the X-chromosome, only one of the two is affected in women only and (2) often a proportion of blood cells belong to the PNH clone. We chose to analyze granulocyte DNA because a large proportion of these cells is usually affected by the PNH abnormality and because they are easy to separate from mononuclear cells. Using this approach, we identified mutations in 12 of 17 PNH patients. HA was more sensitive in localizing frameshift mutations than SSCA and HA. To also pinpoint the mutation in the remaining patients and to further increase the sensitivity of our screening method, we are currently investigating alternative conditions and choices of restriction fragments to be used for both SSCA and HA.

Individual mutations. In 12 patients, sequence analysis of the shifted DNA fragments showed 15 different mutations (Table 1). The mutations were spread throughout the entire coding region without obvious clustering. In exon 2, the largest exon of the PIG-A gene, we found 9 mutations; in addition, 3 were identified in exon 3, 1 in exon 4, and 1 in exon 5.

The mutations identified were all different, and none was identical to the 9 mutations we described earlier.3,5 However, two of the mutations described here have also been reported recently by Yamada et al.5 Both of these are C → T changes (in HH54 the C → T change is on the noncoding strand). This single base pair substitution occurs in CpG dinucleotides and could occur by deamination, to which methylcytidine is vulnerable. Consistent with this dinucleotide being a hot spot for mutation,19 examples of multiple mutation at CpG sites are frequent in mutation databases.

Multiple clones in individual patients. In 3 patients (HH24, HH50, and HH54), we identified two coexisting PNH clones with different mutations. In all 3 patients, the mutations were located on independent M13 clones of the cloned PCR product; therefore, the mutations must have arisen independently. The existence of two different PNH clones in a single PNH patient has now been reported in 6 different patients.4,8 We did not find any obvious clinical pattern that would predict the existence of two independent PNH clones. The only consistent finding was that both mutations identified accounted for a nonfunctional PIG-A protein (PNH III phenotype).

Considering the PNH phenotype of RBCs in our 12 patients, we find that 5 had, in addition to the PNH III, a PNH II RBC population (Table 1). The mutations found in these patients accounted only for the PNH III cells. We can infer that these 5 patients must have an additional PNH clone with a distinct mutation accounting for the PNH II phenotype; therefore, at least 6 of 12 PNH patients have at least two PNH clones. In addition, in patient HH24, in whom we have identified two PNH III mutations at the molecular level, a third PNH clone (PNH II) must also be present. Thus, the majority of our PNH patients have more than one PNH clone that must have arisen independently and coexist with each other. This finding strongly supports the hypothesis of the relative growth or survival advantage of PNH cells in patients with PNH. As to why we have thus far only identified the mutation responsible for the PNH III phenotype in all
patients with both PNH II and PNH III RBCs, the reason is probably that the percentage of PNH II granulocytes in these patients appears to be very small (below the threshold for detection by our method of analysis). On the other hand, PNH II RBCs are readily detected, probably because the residual CD59 expression protects them against complement-mediated lysis and, therefore, increases their life span in circulation.

**PNH versus AA-PNH.** Of our 12 patients, 4 had a history of AA with a negative Ham test before the diagnosis of PNH. A mutation in the PIG-A gene was found in these patients just as it was in patients with classical PNH. Of the 6 mutations identified in these 4 patients, 4 were single base pair substitutions. The number is too small for us to make any firm statement as to the difference in the nature of the mutations in the two groups.

**Mutation pattern in the PIG-A gene compared with that in other genes.** Because the database of PIG-A mutations in PNH has grown to a respectable size, it is now possible to compare its spectrum with that observed in other systems (see Table 2). For this comparison we have selected two tissue-specific genes (β-globin and factor VIII) and one housekeeping gene (glucose-6-phosphate dehydrogenase [G6PD]). Mutations in all of these genes cause relatively common inherited disorders (β-thalassemia, hemophilia A, and G6PD deficiency, respectively).20,21 We have also included the adenomatous poliposis coli (APC) gene, in which both inherited and somatic mutations contribute to the pathogenesis of colon carcinoma.22 The most striking contrast is observed between PIG-A and G6PD. In structural terms, G6PD mutations are virtually all single base pair changes that entail single aa replacements, whereas in the PIG-A gene the majority are insertion-deletion mutations that cause frameshifts. In functional terms, all G6PD mutants have some residual enzyme activity; therefore, it has been suggested that complete inactivation of this enzyme would not be compatible with human life.23 The closest similarity is between the somatic mutations in the APC gene and the somatic mutations in the PIG-A gene.

In conclusion, the spectrum of mutations in the PIG-A gene may reflect the combination of two facts; ie, the gene is housekeeping but its mutations are tissue-specific. We previously suggested that the PNH mutation may confer a relative selective advantage on hematopoietic cells with regard to growth or survival in the context of bone marrow failure.4,24 The predominance of null mutations probably reflects the fact that the total absence of GPI-linked proteins entails a relative survival or growth advantage of the affected cells that is greater than that when the deficiency of GPI-linked proteins is only partial.

**ACKNOWLEDGMENT**

We are grateful to Drs H. Japp (Zurich, Switzerland), W. Gmuer (Zurich, Switzerland), M.-F. McMullin (Belfast, Ireland), T.C. Pearson (London, UK), C. Tsatalas (Thessaloniki, Greece), and C. Bunch and D. Moir (Oxford, UK), all of whom provided us with blood or DNA samples from their patients. We would also like to thank J. Kaeda for his helpful suggestions in SSCA, and D. Stevens for technical assistance (RPMS, London, UK). Also, we are grateful to Dr T. Kinoshita (Osaka University, Osaka, Japan), who provided us with his patient data before their publication.

### Table 2. Frequency of Independently Arising Mutations in Different Genetic Disorders

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Somatic (APC %)</th>
<th>APC (%)</th>
<th>β-Globin (%)</th>
<th>FVIII (%)</th>
<th>G6PD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point mutations</td>
<td>22 (38)</td>
<td>47 (46)</td>
<td>69 (40)</td>
<td>65 (38)</td>
<td>416 (52)</td>
</tr>
<tr>
<td>Nonsense</td>
<td>4</td>
<td>36</td>
<td>57</td>
<td>14</td>
<td>95</td>
</tr>
<tr>
<td>Missense</td>
<td>9</td>
<td>5</td>
<td>6</td>
<td>13</td>
<td>209</td>
</tr>
<tr>
<td>Splice</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>38</td>
<td>12</td>
</tr>
<tr>
<td>Frameshift mutations</td>
<td>35 (60)</td>
<td>56 (54)</td>
<td>105 (60)</td>
<td>46 (27)</td>
<td>49 (7)</td>
</tr>
<tr>
<td>Deletion</td>
<td>14</td>
<td>41</td>
<td>94</td>
<td>26</td>
<td>35</td>
</tr>
<tr>
<td>Insertion</td>
<td>7</td>
<td>15</td>
<td>11</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Deletion-insertion</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>Large deletion</td>
<td>1 (2)</td>
<td>-</td>
<td>30 (18)</td>
<td>83 (10)</td>
<td>-</td>
</tr>
<tr>
<td>Inversion</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>244 (30)</td>
<td>-</td>
</tr>
<tr>
<td>Others</td>
<td>-</td>
<td>-</td>
<td>28 (161)</td>
<td>9 (1)†</td>
<td>2 (3)†</td>
</tr>
<tr>
<td>Total</td>
<td>58</td>
<td>103</td>
<td>174</td>
<td>169</td>
<td>801</td>
</tr>
</tbody>
</table>

**Gene product activity**

| Total deficit             | 47 (81)         | NA     | NA           | 109 (65)§ | 550 (69)  |
| Partial deficit           | 4               | NA     | NA           | 51 (30)§  | 239 (30)  |
| ?†                        | 7               | -      | 28 (95)      | 12 (1)    | -        |

**Abbreviations:** FVIII, factor VIII; NA, not applicable.

*β-Globin mutations include β-, (ββ')-thalassemia and (ββ') HPFH.20,21
†Transcriptional mutants, RNA cleavage/polyadenylation, Cap site.
‡Deletion or insertion in frame.
§For β-thalassemia, total and partial deficit correspond to βA and βB, respectively.
† For FVIII mutation (hemophilia A), total and partial deficit correspond to severe (FVIII < 1%) and moderate to mild (1 < FVIII < 30%) forms, respectively.22
† Unknown amount activity of gene product.
REFERENCES


23. Vulliany T, Beutler E, Luzzatto L: Variants of glucose-6-phosphate dehydrogenase are due to missense mutations spread throughout the coding region of the gene. Hum Mutat 2:159, 1993


Mutations in the PIG-A gene causing paroxysmal nocturnal hemoglobinuria are mainly of the frameshift type

K Nafa, PJ Mason, P Hillmen, L Luzzatto and M Bessler