The EMBO Journal vol.13 no.1 pp.110-117, 1994

Paroxysmal nocturnal haemoglobinuria (PNH) is caused by somatic mutations in the PIG-A gene

Monica Bessler[1], Philip J.Mason[1], Peter Hillmen[1], Toshio Miyata[2], Norio Yamada[3], Junji Takeda[2], Lucio Luzzatto[1,4] and Taroh Kinoshita[2]

[1]Department of Haematology, Hammersmith Hospital, Royal Postgraduate Medical School, Du Cane Road, London W12 0NN, UK, [2]Department of Immunoregulation, Research Institute for Microbial Diseases, Osaka University, Suita, Osaka 565 and [3]Department of Internal Medicine, The Branch Hospital, Nagoya University School of Medicine, Nagoya 461, Japan
[4]Corresponding author

Communicated by L.Luzzatto

Paroxysmal nocturnal haemoglobinuria (PNH), an acquired clonal blood disorder, is caused by the absence of glycosyl phosphatidylinositol (GPI)-anchored surface proteins due to a defect in a specific step of GPI-anchor synthesis. The cDNA of the X-linked gene, PIG-A, which encodes a protein required for this step has recently been isolated. We have carried out a molecular and functional analysis of the PIG-A gene in four cell lines deficient in GPI-linked proteins, obtained by Epstein-Barr virus (EBV) transformation of affected B-lymphocytes from PNH patients. In all four cell lines transfection with PIG-A cDNA restored normal expression of GPI-linked proteins. In three of the four cell lines the primary lesion is a frameshift mutation. In two of these there is a reduction in the amount of full-length mRNA. The fourth cell line contains a missense mutation in PIG-A. In each case the mutation was present in the affected granulocytes from peripheral blood of the patients, but not in normal sister cell lines from the same patient. These data prove that PNH is caused in most patients by a single mutation in the PIG-A gene. The nature of the mutation can vary and most likely occurs on the active X-chromosome in an early haematopoietic stem cell.

Key words: paroxysmal nocturnal haemoglobinuria/PIG-A/mRNA stability/somatic mutation

Introduction

Paroxysmal nocturnal haemoglobinuria (PNH) is an acquired haemolytic blood disorder often associated with cytopenia and recurrent venous thrombosis (Dacie and Lewis, 1972; Rotoli and Luzzatto, 1989). Previous evidence indicates that PNH results from the expansion of an abnormal clone which arises by a somatic mutation occurring in an early haematopoietic stem cell (Oni *et al.*, 1970). Cells of the abnormal clone are deficient in all cell surface proteins which are normally anchored to the membrane by a glycosyl phosphatidylinositol (GPI) molecule (De Sandre *et al.*, 1956; Nicholson-Weller *et al.*, 1983; Pangburn *et al.*, 1983; Hansch *et al.*, 1987; Selvaraj *et al.*, 1988; Rosse *et al.*, 1990). At least two of these missing proteins protect blood cells against complement attack. The deficiency in GPI-linked proteins is due to a defect in the GPI-anchor assembly in the endoplasmic reticulum. The defect has been pinpointed to the step catalysed by the α-1,6-*N*-acetylglucosaminyl-transferase (Armstrong *et al.*, 1992; Hillmen *et al.*, 1993c; Takahashi *et al.*, 1993). Recently, the cDNA corresponding to the gene responsible for this particular step in the anchor biosynthesis, named PIG-A (for phosphatidylinositol glycan complementation class A), has been isolated by expression cloning, sequenced and mapped to the X-chromosome (Miyata *et al.*, 1993; Takeda *et al.*, 1993). The PIG-A cDNA contains a 1452 bp open reading frame that encodes a putative membrane protein of 484 amino acid residues (Miyata *et al.*, 1993). Whether the PIG-A gene actually encodes the α-1,6-*N*-acetylglucosaminyltransferase is not yet certain, but in two Japanese patients with PNH it has been shown that an abnormality in the PIG-A gene is responsible for the disease. In one patient a single base pair deletion at a 5′ splice site caused abnormal splicing and a truncated mRNA. In the other patient a reduced amount of normal sized RNA was found (Takeda *et al.*, 1993).

In the present study we have analysed four cell lines deficient in GPI-anchored proteins which we have obtained by Epstein-Barr (EBV) transformation of the affected B cells from four British patients with PNH (Hillmen *et al.*, 1993b). Each one of the four cell lines has an abnormality in the PIG-A gene. In three cell lines we found frameshift mutations and two of these are associated with a severely reduced level of normal sized RNA. The fourth cell line had a point mutation in the coding region of the PIG-A gene. We have verified that these lesions are due to somatic mutations because they are not present in normal cell lines isolated from the same patients. These data confirm that mutations of the PIG-A gene are responsible for PNH in at least the majority of patients, and they show that a whole array of different mutations in the same gene can cause the PNH phenotype.

Results

Correction of the PNH phenotype by transfection of the PIG-A cDNA

We have previously obtained from patients with PNH EBV-transformed lymphoblastoid cell lines which have been fully characterized and shown to be deficient in GPI-linked proteins (Hillmen *et al.*, 1993b). Four such PNH cell lines from four different patients were transfected by electroporation with a plasmid competent for expression of PIG-A cDNA and then tested for the presence of surface CD59. Expression of this molecule was promptly restored in each case (Figure 1). When cells were grown in a hygromycin-containing medium after transfection, there was gradual enrichment for CD59-positive cells. After 20 days practically all cells had an almost normal amount of CD59 (Figure 1). The transfected cells expressing CD59 also expressed CD55

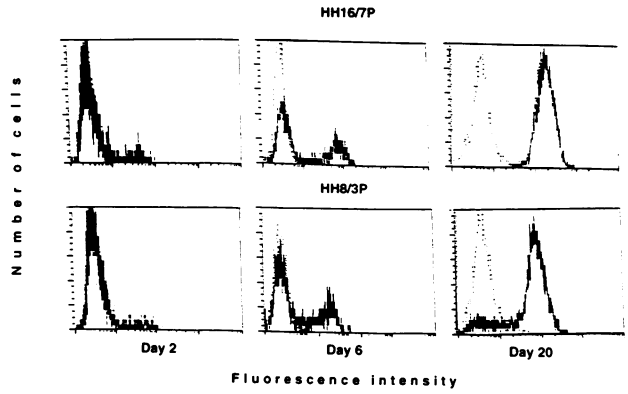

Fig. 1. CD59 expression on the surface of PNH cells after transfection with PIG-A cDNA. The two PNH cell lines HH16/7P and HH8/3P are derived from two different PNH patients. The cells were transfected with the PIG-A cDNA (continuous line) or an irrelevant control cDNA (CD8) (dotted line) on day 0. Selection for the transfectants in hygromycin-containing medium was started on day 2. CD59 expression was determined by flow cytometry using an anti-CD59 monoclonal antibody.

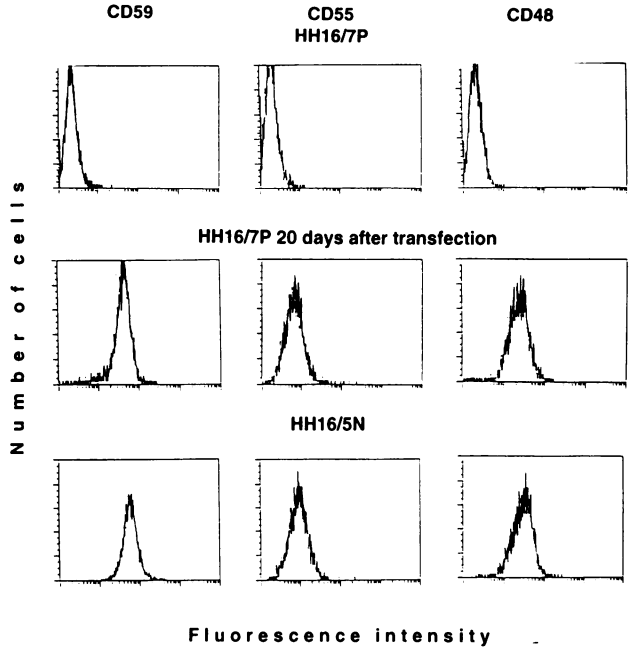

Fig. 2. Expression of several GPI-linked proteins on the surface of PNH cells before and after transfection with PIG-A cDNA and selection in hygromycin. CD59, CD55 (DAF) and CD48 on the surface of the PNH cells were determined 20 days after transfection using the appropriate monoclonal antibody and flow cytometry. Top row, untransfected cell line HH16/7P; middle row, PNH cell line HH16/7P 20 days after transfection; bottom row, the phenotypically normal sister cell line HH16/5N.

and CD48 at a level comparable with that of normal lymphoblastoid cell lines (see Figure 2).

Analysis of PIG-A transcripts in PNH cell lines

Since PIG-A fully corrects the PNH phenotype, we were satisfied that it must have been the defective molecule in these cell lines. Specifically, it seemed likely that a mutation in the PIG-A gene is responsible for the PNH phenotype. Poly(A)-enriched mRNA from each of the four cell lines was analysed by Northern blotting. For comparison we used HeLa cell RNA, RNA from a normal lymphoblastoid cell line obtained from a normal individual, as well as RNA from

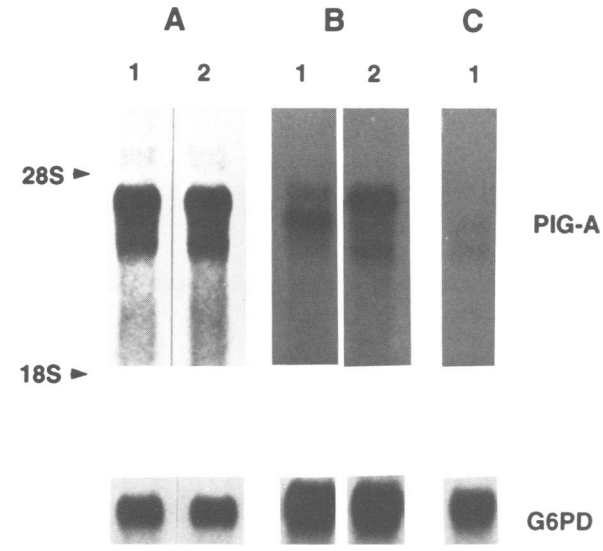

Fig. 3. Northern blot analysis of PIG-A mRNA in normal and PNH lymphoblastoid cell lines. Ten micrograms of poly(A)-enriched RNA were applied to each lane. Filters were hybridized with the PIG-A cDNA probe. The films were exposed at −80°C for 14 days. The same filters were subsequently hybridized to a G6PD cDNA probe (Persico *et al.*, 1986) and exposed overnight. Lanes 1A, 1B and 1C: mRNA from the PNH cell lines HH16/7P, HH8/3P and HH5/7P. Lanes 2A and 2B: RNA from the corresponding phenotypically normal sister cell lines. (No phenotypically normal cell line was available from patient HH5.) The position of 18S and 28S rRNA bands is shown on the left.

normal lymphoblastoid cell lines obtained from three of the four patients from whom the PNH cell lines were derived.

In normal cells we found three different mRNA species. Based on the known sizes of 28S (Conzalez *et al.*, 1985) and 18S RNAs (McCallum *et al.*, 1985), we estimate their sizes to be 3.8, 3.3 and 2.8 kb, respectively (see Figure 3). In the PNH cell lines HH20/9P and HH16/7P, the mRNA patterns were indistinguishable from those of their normal sister cell lines HH20/17N and HH16/5N or of lymphoblastoid cells from a normal person and of HeLa cells. On the other hand, the PNH cell line HH8/3P showed a sharply different pattern from that of its normal sister cell line HH8/4N. Although all three mRNA species were present, their relative abundance was markedly altered. Specifically, the medium-size (3.3 kb) mRNA, hardly visible in the normal cell line, was the major mRNA species in the PNH cell line; by contrast, the largest (3.8 kb) was barely visible (Figure 3). The PNH cell line HH5/7P (for which we do not have a normal sister cell line) has yet a different pattern: here the 3.8 kb mRNA is missing altogether, whereas the two smaller mRNA species are present in apparently normal amounts (Figure 3).

In order to further define the abnormalities of PIG-A mRNA, this was reverse transcribed and the coding region was amplified by polymerase chain reaction (PCR). From RNA of two normal lymphoblastoid cell lines obtained from normal persons and from HeLa cells, as well as from all three phenotypically normal lymphoblastoid cell lines obtained from three of our PNH patients, three major amplification products were detected using primers 11 and 12. The largest amplification product has the predicted size of the coding region (1500 bp); two additional products were ~1200 and 800 bp in size, respectively (Figure 4, lane 3). Sequence analysis showed that the largest band was the

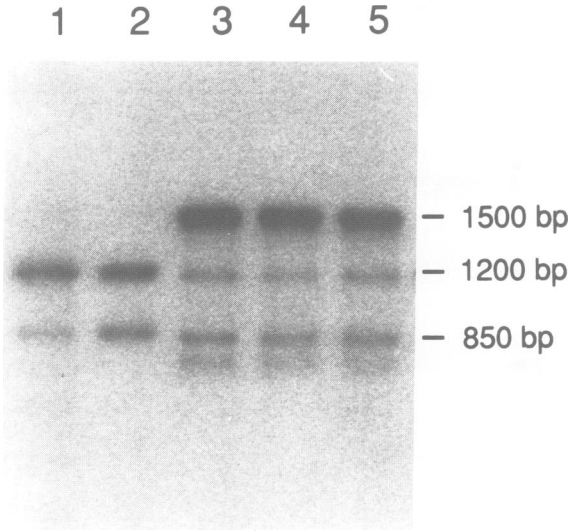

Fig. 4. Autoradiograph of RT−PCR analysis of PIG-A mRNA from PNH cell lines. The entire coding region of the PIG-A cDNA was amplified using primers 11 and 12 (Table I). The radiolabelled RT−PCR products were analysed on a 2% agarose gel. Lane 1, HH8/3P; lane 2, HH5/7P; lane 3, JY cell; lane 4, HH20/9P; lane 5, HH16/7P.

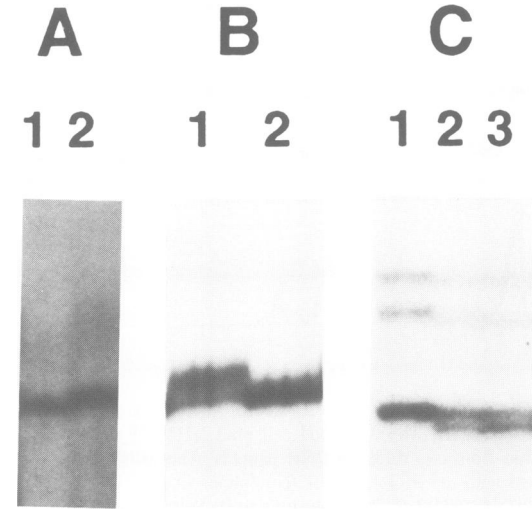

Fig. 5. SSCP analysis used to localize mutations found in PNH cell lines. (**A**) Heat-denatured 180 bp *Fok*I restriction fragment from the 3′ amplification product of the PIG-A cDNA (primers c and f, Table I) from cell lines of patient HH20. Lane 1, HH20/9P; lane 2, HH20/17N. (**B**) Non-denatured 270 bp *Hin*dIII−*Taq*I fragment of exon 2 amplified from genomic DNA (primers −I1 and I2, Table I) from lymphoblastoid cell lines of patient HH8. Lane 1, HH8/3P; lane 2, HH8/4N. (**C**) Non-denatured 340 bp *Hin*dIII−*Taq*I fragment of exon 2 amplified from genomic DNA of PNH cell line HH5/7P (lane 1), and from the patient's granulocytes (lane 2) and lymphocytes (lane 3). For patient HH20 (A), the mutation was localized by the band shift present in the denatured fragments after digestion with *Fok*I (A) and *Bst*NI (not shown). In patient HH8, who is a female PNH patient and who turned out to be heterozygous for the mutation, the two bands for the non-denatured 270 bp *Hin*dIII−*Taq*I fragment are due to different electrophoretic mobility of the heteroduplex. Heteroduplex formation is also the cause of the two bands in the non-denatured 340 bp *Hin*dIII−*Taq*I fragment from granulocytes and lymphocytes (lanes 2C and 3C) of patient HH5 (C), who is a male patient and has a mixture of normal and PNH cells in his peripheral blood.

product to be expected from the primers used, i.e. it had the sequence of PIG-A cDNA from nucleotide −36 to 1476 (numbers refer to cDNA unless stated otherwise). The second largest band had the same sequence, except for a 374 bp excision from nucleotide 342 to 715. The third band had a 658 bp excision from nucleotide 58 to 715. Analysis of the genomic structure revealed that exon 2 of the PIG-A gene extends from nucleotide −62 to nucleotide 715 (data not shown). At each of the two positions 58 and 342 we note a GT dinucleotide, which is a potential 5′ splicing site. We presume, therefore, that the three reverse transcription (RT)−PCR products reflect mRNA species arising by several alternative splicing pathways, in which portions of exon 2 are spliced out of the primary transcript. Thus, it seems likely that only the 1500 bp RT−PCR product corresponds to functional mRNA. It appears that normal cells also produce smaller, presumably non-functional mRNA species.

When RNA from our PNH cell lines was similarly subjected to RT−PCR analysis, HH20/9P and HH16/7P yielded patterns which were indistinguishable in size and relative amounts from the products of normal lymphoid cells (Figure 4, lanes 4 and 5). In contrast, RNA from cell lines HH8/3P and HH5/7P yielded only a major band of ~1200 bp and a minor band of ~800 bp, but no band of 1500 bp, corresponding to the full size of the mRNA coding region (Figure 4, lanes 1 and 2).

In order to assess the functional competence of the PIG-A mRNA from PNH cell lines, the largest cDNA products obtained by RT−PCR were cloned into the pEB expression vector and transfected into the GPI-anchor-deficient mutant cell line JY5 (originally used for expression cloning of the PIG-A cDNA; Miyata *et al.*, 1993). Transfection of the 1500 bp RT−PCR product from the JY cell line (the phenotypically normal parent cell line of JY5) restored expression of the surface antigen CD59 on JY5. By contrast, transfection of the 1500 bp product of HH20/9P and HH16/7P failed to restore CD59 expression, indicating that the apparently normal-sized mRNAs from these two cell lines are non-functional, and are therefore likely to have a mutation in the coding region. The 1200 bp cDNA products of HH8/3P and HH5/7P also failed, not surprisingly, to restore CD59 expression after transfection into the JY5 cell line.

Analysis of the molecular lesions in PNH cell lines at the DNA level

For the two cell lines HH20/9P and 16/7P, the mRNA patterns combined with the results of the transfection experiments suggested point mutations within the coding region. SSCP analysis of HH20/9P cDNA and comparison with the analysis of HH20/17N showed a band shift close to the 3′ end of the coding region. *Fok*I and *Bst*NI digests of the 700 bp sized RT−PCR amplification products localized the mutation between nucleotides 1331 and 1388. By sequence analysis we found that indeed in HH20/9P an AT doublet at position 1348−1349 is replaced by a single C.

The mutation in HH16/7P could not be located by SSCP analysis, but direct sequencing of the PCR product of exon 2 revealed a C to T replacement at position 464.

For cell lines HH8/3P and HH5/7P, which had altered mRNA patterns, SSCP analysis of *Hin*dIII/*Taq*I digests and of *Bst*XI digests of exon 2 amplified from genomic DNA showed band shifts which localized the mutation between nucleotides 228 and 425 for cell line HH8/3P, and between

nucleotides 505 and the 3′ end of the 996 bp amplification product for cell line HH5/7P (see Figure 5). By sequencing the entire exon 2 we found indeed that in HH8/3P a GC doublet at positions 336–337 was replaced by a single T (Figure 6A). In HH5/7P we found instead a 5 bp insertion, TAGAT, at position 634. Since the primers used (−I1 and I2; Table I) are located 54 bp 5′ and 115 bp 3′, respectively, of the intron–exon boundaries of exon 2 we were able to verify that these flanking regions, including the splice sites, had an entirely normal sequence.

The PNH phenotype results from somatic mutations

In order to confirm that the mutations observed are not constitutional, we took advantage of the normal sister cell lines and of fresh patient material when available. In all three patients for whom we did obtain a PNH cell line and a phenotypically normal cell line, the mutation present in the PNH cell line was not present in the normal sister cell line. For patient HH5 for whom no phenotypically normal cell lines were available, both normal and mutated genes were shown to be present in DNA from the patient's mononuclear cells by sequencing M13 clones. HH8 was a female patient, and the cell line HH8/3P was found by sequence analysis to be heterozygous for the mutation described. This mutation predicts loss of a *Dde*I restriction site, and this was shown to be missing in the PNH cell line HH8/3P (thus also ruling out a possible PCR artefact), whereas it was retained in the patient's normal cell line HH8/4N (see Figure 7). By using the same approach, we have been able to show that all of the cDNA from the PNH cells lacks the *Dde*I site (see Figure 7), indicating that there is no detectable transcript from the normal allele.

We next confirmed the presence of the mutation in the patient's peripheral blood granulocytes. When the amplification product of exon 2 from granulocyte genomic DNA was cloned into M13, two out of 14 clones sequenced showed the nucleotide change at position 336, whereas the remaining 12 were normal. Considering that the PNH granulocytes are also heterozygous, the proportion of M13 clones with the mutation (~14%) fits reasonably well with

Table I. Oligonucleotides used in reverse transcription and PCR analysis of the PIG-A gene

Purpose	Primer sequence (5′→3′)	Designation	Position (5′→3′)	Size of the amplified product (bp)
Primer for cDNA synthesis	AATGATATAGAGGTAGCATAAC	13	1586–1565	
Amplification of portions of cDNA	ACTGGCGGCCATGGAACTCAC[a]	a	−85 to −65	641
	AAGACACACAAATGATGTGG	d	556–537	
	TTACAAACAAGCTTCTAACCGT[a]	b	497–518	518
	CTCAGGAATTCCACCAACTCT	e	1014–994	
	TCATTTTGGAAGAAGTTCG[a]	c	806–824	700
	CTTTcTAGAACTATTACAAATGTT	f	1505–1482	
	accagagctcGGTTGCTCTAAGAACTGATGTC	11	−36 to −14	1532
	accaggtaccTCTTACAATCTAGGCTTCCTTC	12	1476–1455	
Amplification of genomic exon 2	AATATCTAgATCTGGATTTGTA[a]	−I1	I1 (−75 to −54)[b]	991
	gggaaTTCTTTGCTATAATTTATAG	I2	I2 (115–134)[c]	

Lower case indicates non-PIG-A sequences designed to introduce restriction enzyme sites.
[a]5′ biotinylated.
[b]Number of bp 5′ of the intron 1–exon 2 boundary.
[c]Number of bp 3′ of the exon 2–intron 2 boundary.

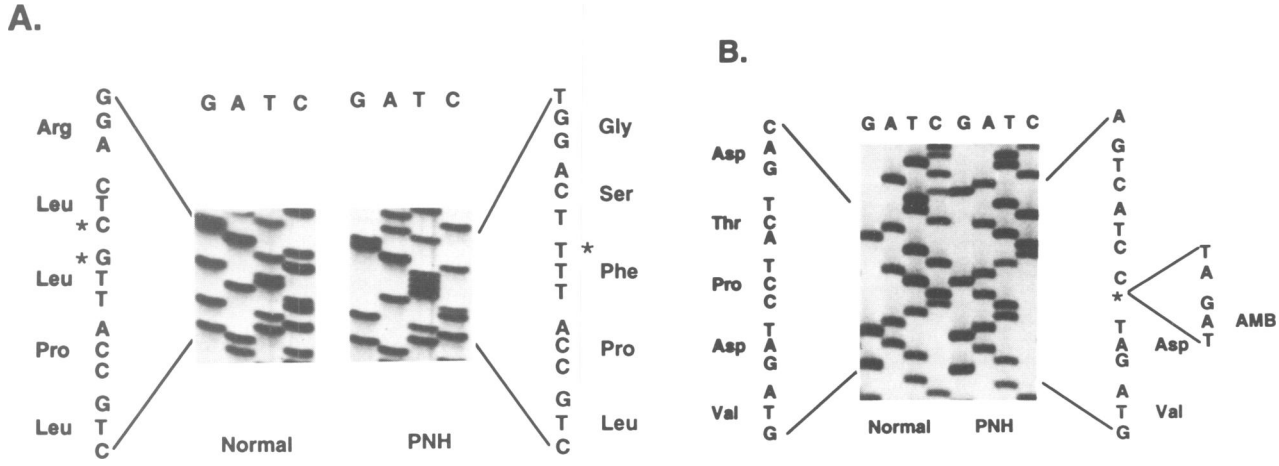

Fig. 6. Nucleotide sequence of genomic DNA encompassing the mutation found in the two PNH cell lines. **(A)** Nucleotide sequence obtained from M13 clones carrying the normal and the mutant allele in HH8/3P. **(B)** Nucleotide sequence of M13 clones obtained from HH5/7P and from the patient's normal lymphocytes.

the proportion of granulocytes deficient in GPI-anchored membrane proteins, previously determined to be 40% in this patient (Bessler *et al.*, 1992). A similar estimate was also obtained by visual inspection of an ethidium bromide-stained agarose gel after electrophoresis of the PCR product digested with *Dde*I (see Figure 7).

In the male patient HH5, the mutation detected in the cell line HH5/7P was also demonstrated in the patient's granulocytes (95% of which belong to the PNH clone) (Figure 6B), and in three out of six M13 clones obtained from the patient's mononuclear cells, again in good agreement with the previous finding that 55% of mononuclear cells were deficient in GPI-anchored membrane proteins.

Discussion

Expression of PIG-A regularly corrects the PNH phenotype

Transfection of our four PNH cell lines with the PIG-A cDNA restored the expression of three different GPI-linked surface proteins which were missing from these cells (Hillmen *et al.*, 1993b). It seems safe to assume that the same applies to all GPI-linked proteins, and that we have thus achieved a full correction of the PNH phenotype. We had previously shown that the absence of GPI-linked surface proteins in all of these cell lines was due to a specific block in GPI-anchor biosynthesis (Hillmen *et al.*, 1993c). The restoration of normal expression of all GPI-linked proteins we have tested thus proves that the product of the PIG-A cDNA is indeed able to relieve the block in GPI-anchor biosynthesis. Along with the data recently reported on other cell lines from two Japanese patients (Takeda *et al.*, 1993), this brings to six the total number of cases in which it has been formally proved that the absence of a normal PIG-A gene product produces the PNH phenotype.

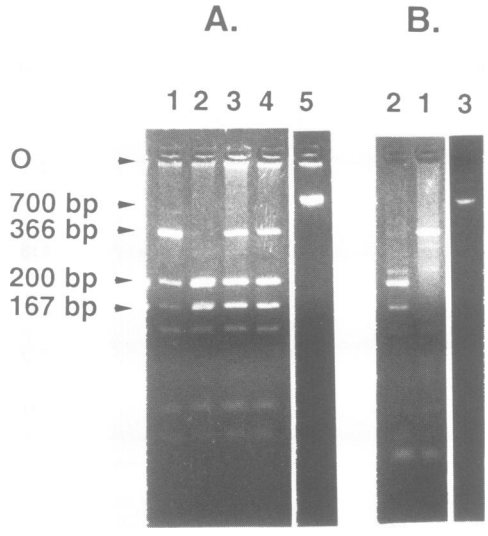

Fig. 7. Analysis of the PNH mutation in genomic DNA and in cDNA of patient material. All samples were from patient HH8. Genomic DNA was amplified using the primers $-I_1$ and d (Table I), which results in a 700 bp product (panel A, lane 5). cDNA was amplified with primers a and d (Table I), yielding a 622 bp product (panel B, lane 3). *Dde*I digestion of these products yields several fragments, the two largest being 200 and 167 bp, respectively (lane 2 in both panels A and B). Since in this patient the PNH mutation abolishes a *Dde*I site, these two fragments are replaced by a single 366 bp fragment. (**A**) Genomic DNA digested with *Dde*I, run on a 3% agarose gel and stained with ethidium bromide. Lane 1, HH8/3P (PNH cell line); lane 2, HH8/4N (normal cell line); lane 3, granulocytes; lane 4, lymphocytes; lane 5, undigested PCR product of HH8/4N. (**B**) RT–PCR products digested with *Dde*I. Lane 1, HH8/3P; lane 2, HH8/4N; lane 3, undigested RT–PCR product of HH8/4N. Analysis of the genomic DNA shows that in HH8/3P undigested and digested fragments are present (panel A, lane 1), in keeping with the cell line being heterozygous for the mutation. In granulocytes (40% PNH cells: panel A, lane 3) and lymphocytes (20% PNH cells: panel A, lane 4) the relative amounts of the restriction fragments are in agreement with the notion that the patient's PNH blood cells are heterozygous for the same mutation. Analysis of the cDNA from HH8/3P with *Dde*I shows the presence of only the undigested 366 bp fragment (panel B, lane 1), demonstrating that only the mutant allele is expressed in the PNH cell line, which is in agreement with the mutation having occurred on the active X chromosome.

Multiple mRNA species from the PIG-A gene

In order to characterize the genetic changes underlying the PNH phenotype, we first chose to analyse the PIG-A mRNA in normal cells. Northern blots of poly(A)-enriched RNA from both phenotypically normal lymphoblastoid cells and HeLa cells revealed three different mRNA species. The most prominent band was estimated to be 3.8 kb in size based on ribosomal markers (Conzalez *et al.*, 1985; McCallum *et al.*, 1985). In addition, we observed consistently a less intense band at 2.8 kb and a very faint band at 3.3 kb. [The smaller mRNA species had not been detected on Northern blots of total RNA (Miyata *et al.*, 1993; Takeda *et al.*, 1993), presumably because of their low abundance.] From comparison with the signal obtained with a housekeeping gene probe (G6PD; see Figure 3), we infer that the PIG-A mRNA has very low abundance in lymphoblastoid cells.

Although we have not formally established the correspondence between these mRNA species and the RT–PCR products, the latter technique confirms the presence of multiple mRNA species. The two smaller PCR products indicate the presence of mRNA molecules in which portions of exon 2 have been eliminated: these must arise through the use of two alternative splice sites within exon 2. From sequence analysis, we find that each of these two mRNA species has a frameshift, with the open reading frame being terminated by a stop codon after only four and seven codons, respectively.

Alternative splicing is often used to produce multiple distinct mRNAs from a single gene, and this may introduce functional diversity amongst the products of a single gene (Smith *et al.*, 1989; Green, 1991). This process may be regulated in a developmentally specific manner or in response to specific physiological conditions to produce stage- or condition-specific mRNAs. In some cases, when the product is not functional, alternative splicing can act as an on/off switch of a gene at a post-transcriptional level: for instance, this has been well documented for a number of *Drosophila* genes (Bingham *et al.*, 1988). In the case of the PIG-A gene, alternative splicing gives rise to several non-functional products, but whether this phenomenon has a regulatory role remains to be determined.

Heterogeneity of PIG-A mutations in PNH

Analysis of the RNA and DNA from our four PNH cell lines has revealed four different mutations, all in the PIG-A gene, causing severely diminished expression of GPI-anchored membrane proteins on the cell surface. Although in two cell lines (HH20/9P and HH16/7P) the RNA pattern was normal, cDNA from all four, inserted in an expression vector, failed to restore the synthesis of GPI-anchored protein in the test

cell line JY5. SSCP analysis of the amplification products obtained from cDNA and genomic DNA, followed by digestion with appropriately selected restriction enzymes, showed a band shift in three cell lines and, therefore, was a useful tool in localizing the site of the mutation in each.

In patient HH20, once the position of the mutation was localized by SSCP analysis, sequence analysis showed that an AT doublet at position 1348–1349 was replaced by a single C. The mutation results in a frameshift which causes a termination of the open reading frame after four codons. The resulting protein would be truncated at residue 454, thus lacking the last 30 C-terminal amino acids. The fact that the PNH cell line is completely deficient in GPI-anchored membrane proteins suggests that this putative protein is not functional. This assumption is supported by the transfection experiment which failed to restore the expression of GPI-anchored membrane proteins in JY5 after transfection of the PIG-A cDNA obtained from this patient's cell line. The loss of function might be due to changes in the active site. An alternative possibility is that the truncated protein is not retained in the membrane of the endoplasmic reticulum and, therefore, is not available where it is needed for GPI-anchor synthesis.

In patient HH16, direct sequencing of exon 2 revealed a C to T change at position 464, resulting in the substitution of phenylalanine for serine at amino acid position 155. The fact that the PNH phenotype does indeed result from this amino acid replacement was proven by its absence in the normal sister cell line and by the functional transfection assay. Presumably, this single amino acid replacement causes functional inactivation of PIG-A.

The PNH cell lines of patients HH8 and HH5 showed abnormal mRNA patterns by Northern blot and RT–PCR analysis. In HH8/3P, we found that a GC doublet was replaced by a single T at position 336. Although the mutation is only 6 bp 5′ to the alternative splice at position 342, it does not affect the 5′ consensus (Oshima and Gotoh, 1987). It does, however, produce a frameshift and termination of the open reading frame after 11 codons in the full-length PIG-A mRNA. On the other hand, the mutation produces an open reading frame of 1080 bp in the second largest transcript. Since the PNH cell line lacked the expression of GPI-anchored proteins and no expression of CD59 was achieved after transfection of the 1200 bp RT–PCR product into the JY5 cells, the putative protein of 360 amino acids cannot have any activity with respect to GPI-anchor synthesis.

In the PNH cell line of patient HH5, a 5 bp insertion (TAGAT) at position 634 also produces a frameshift and inserts an in-frame stop codon. The insertion is 81 nucleotides upstream of the 3′ end of exon 2, and therefore is unlikely to affect splicing. The 1200 and 800 bp RT–PCR products, the only ones found in this cell line, are unaffected by the mutation since the 3′ end of exon 2, which includes position 634, is spliced out in both of the smaller products.

In patient HH8, the mutation we have identified might alter splicing because it is very close to the splicing site, and altered efficiency of splicing by a mutation not very close to a splice site (as in patient HH5) has been reported (Smith *et al.*, 1989). However, an alternative and perhaps more attractive explanation of the abnormal RNA patterns observed in these two cell lines is that premature termination of translation decreases the stability of the PIG-A mRNA (Urlaub *et al.*, 1989). This would explain why the two smaller transcripts of normal lymphoid cell lines, which have stop codons, have low abundance; at the same time, it would explain the reduced amount or complete lack of the full-length transcript in HH8/3P and HH5/7P. Low abundance of a mutated RNA due to an early stop codon has been well documented for β-globin mRNA in certain cases of thalassaemia (Baserga and Benz, 1988). By the same mechanism, restoration of the reading frame in the mRNA corresponding to the second RT–PCR product in HH8/3P would explain the prominent middle-sized mRNA species seen in the Northern blot (Figure 3). It is interesting that, by contrast, the premature termination of transcription did not markedly alter the relative abundance of the largest sized RNA species in patient HH20, presumably because here the abnormal stop codon is near the normal stop codon. Thus, analysis of the molecular defect in these four PNH patients has revealed in each a different mutation within the coding region of the PIG-A gene, and all of these mutations are different from the previously reported mutation found in a Japanese patient, which was in a 5′ splice site and prevented production of full-length mRNA.

Somatic mutations in the PIG-A gene are responsible for PNH in all patients tested thus far

With the single exception of the patient suffering from a similar condition as a result of an inherited CD59 deficiency, in all patients with PNH the mutation is always in the PIG-A gene. The fact that the genetic lesions found in PNH are heterogeneous is not surprising, since we are dealing with somatic rather than inherited mutations. The most likely explanation for the mutation being always in the PIG-A gene is that, as previously suggested by somatic cell hybridization experiments (Hillmen *et al.*, 1993a), and proven by direct mapping (Takeda *et al.*, 1993), the PIG-A gene is on the X chromosome. Thus, deficiency of other enzymes involved in the GPI-anchor biosynthetic pathway, which are probably encoded in autosomal genes, would require inactivating mutations of both alleles, whereas with PIG-A just one inactivating mutation is sufficient. Whereas with inherited mutations X-linkage is classically associated with a much higher frequency of disease in males than in females, this is not true in the case of PNH. The explanation is clearly that at the somatic cell level, as a result of X chromosome inactivation, there is only one active X in each cell in both males and females. The PIG-A gene maps on the short arm of the X, but not within the pseudo-autosomal region, and the data reported in two cases (Takeda *et al.*, 1993 and patient HH8, see Figure 7) indicate that PIG-A is subject to inactivation in female somatic cells. Thus, the PNH mutation must have taken place on the active X chromosome after X chromosome inactivation, and therefore it is responsible for the disease in these patients.

Previous evidence suggested that PNH is due to the expansion of an abnormal cell clone thought to arise by a somatic mutation in an early haematopoietic stem cell (Oni *et al.*, 1970; Dacie and Lewis, 1972; Rotoli *et al.*, 1984). More recent biochemical studies suggested that a single metabolic block might be responsible for PNH in the majority of patients (Armstrong *et al.*, 1992; Hillmen *et al.*, 1993c; Takahashi *et al.*, 1993), consistent with complementation analysis using somatic cell fusion between GPI-anchor-deficient mutant mouse thymoma cell lines (Thy-1

mutants) and lymphoblastoid cell lines obtained from PNH patients (Armstrong *et al.*, 1992; Takahashi *et al.*, 1993). Somatic cell hybridization between normal and PNH lymphoblastoid cell lines showed that the mutant allele was recessive (Hillmen *et al.*, 1993a). This work, together with the previous work by Takeda *et al.* (1993), now proves directly the somatic mutation theory of the pathogenesis of PNH, because the mutations in the PIG-A gene found in PNH cell lines are not present in normal sister cell lines from the same patients or in the normal cells from the patient's peripheral blood.

Materials and methods

Cell lines

The production of the B-lymphoblastoid cell lines from the PNH patients (normal: HH20/17N, HH16/5N, and HH8/4N; PNH HH20/9P, HH16/7P, HH5/7P and HH8/3P) has been described previously (Hillmen *et al.*, 1993b). The lymphoblastoid cell line JY and its GPI-deficient mutant JY5 (Holander *et al.*, 1988) were also used. The cell lines were cultured in RPMI 1640 medium containing 10% fetal calf serum.

Transfection with PIG-A cDNA

Full-length cDNA was cloned into the EBV-based mammalian expression vector pEB (Miyata *et al.*, 1993). The PIG-A cDNAs prepared by RT–PCR (see below) were also cloned into the pEB vector. Ten micrograms of the pEB containing the PIG-A cDNA were transfected into $1-3 \times 10^7$ lymphoblastoid cells by electroporation. The electroporation was carried out with a Gene-PulserTM₁ (Bio-Rad) at 220–250 V and 960 μFD in 0.8 ml culture medium. To select for transfectants, 50–200 μg/ml hygromycin B (Calbiochem) were added to the culture medium 2 days after transfection. Dead cells were removed repetitively by centrifugation over a Ficoll gradient. Transfection conditions and hygromycin concentration varied for each cell line and were optimized previously. A cDNA encoding CD8, an antigen usually not present on B-cells, in a similar vector [EBO-pcD-Leu2 (Margolskee *et al.*, 1988) generously provided by R.F.Margolskee, Roche Institute of Molecular Biology, Nutley, NJ] was used as a control plasmid.

Analysis of GPI-anchored membrane proteins

Quantitation of the expression of GPI-linked membrane proteins was performed by flow cytometry (FACscan, Becton Dickinson) using monoclonal antibodies towards CD59 (Bric229, Bioproducts Laboratories), CD55 (IA10, a gift from Prof. V.Nussenzwig, Department of Pathology, NYU Medical Center, New York, NY), CD48 (LoMN25, a gift from Dr B.De La Parra, University of Louvain, Louvain, Belgium) as described previously (Hillmen *et al.*, 1993b).

Analysis of PIG-A mRNA

Total RNA was isolated using the guanidine isothiocyanate method (Chirgwin *et al.*, 1979). Poly(A)$^+$ RNA was obtained by oligo dT-cellulose column chromatography (Collaborative Research Incorporated). Northern blots were prepared using formaldehyde-containing gels (Mason *et al.*, 1993). The blots were probed with the PIG-A cDNA probe. This probe corresponds to nucleotides −76 to 2606 of the PIG-A cDNA (Takeda *et al.*, 1993).

Generation of cDNA-PCR products for functional analysis and gel analysis

Total RNA (4 μg) was reverse transcribed using 50 pmol of primer 13 (Table I) with 200 U of RNase H-free reverse transcriptase (SuperscriptTM₁ Gibco BRL) in 50 mM Tris–HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 100 μg/ml bovine serum albumin (BSA) in a total volume of 20 μl. The coding region of PIG-A was amplified from the cDNA with 50 pmol of each primer 11 and 12 (Table I) in 50 μl of 10 mM Tris–Cl (pH 8.3), 1.5 mM $MgCl_2$, 0.001% gelatine and 25 μM dNTP by 25 cycles of a reaction consisting of incubation for 1 min at 93°C, 1 min at 55°C and 3 min at 72°C; 10 μCi [α-^{32}P]dCTP (Amersham) were used to radiolabel PCR products. The PCR products were analysed by electrophoresis in 2% agarose gel, followed by autoradiography.

Generation of cDNA-PCR products for SSCP analysis and sequencing

Total RNA (2 μg) was used in the reverse transcriptase reaction. Reverse transcription was performed for 2 h at 37°C in 50 mM Tris–HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DDT, 1 mM each of dATP, dCTP, dTTP and dGTP, and 105 μg/ml random hexamer primer (pd$(N)_6$, Pharmacia) in a total volume of 40 μl using 300 U of reverse transcriptase M-MLV RT (Gibco BRL) and 30 U of ribonuclease inhibitor RNasin (Promega). The reaction was terminated by incubation at 65°C for 10 min and centrifugation for 2 min at 4°C. Three microlitres of cDNA were used for each 100 μl PCR reaction.

DNA amplification

DNA was prepared by the SDS/Proteinase K extraction procedure (Sambrook *et al.*, 1989). For the amplification of exon 2 from genomic DNA, two intronic primers were chosen (Table I). The primers used to amplify the 1452 bp coding region of the PIG-A cDNA are shown in Table I. The PCR was performed in a total volume of 100 μl containing 0.5 mM dNTPs, 100 ng of each primer, 2 U of Taq DNA polymerase (Boehringer Mannheim), 16.6 mM $(NH_4)_2SO_4$, 67 mM Tris–Cl (pH 8.8), 6.7 mM $MgCl_2$, 10 μg gelatine and 10 mM β-mercaptoethanol. The amplification reaction was carried out using a Perkin Elmer Cetus thermal cycler (Perkin Elmer Cetus Instruments). A 30 cycle protocol was used with denaturation at 95°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 3 min. Reactions were electrophoresed in a 1–1.5% agarose gel, stained in ethidium bromide and visualized on a UV transluminator.

Analysis by single-strand conformation polymorphism

The entire exon 2 amplified from genomic DNA, including the flanking intron–exon boundaries and the entire coding region of the PIG-A cDNA amplified in three portions using the appropriate cDNA as template (Table I), were screened for mutation by SSCP analysis. Fragments were labelled during the amplification reaction by adding 0.1 μl of [α-^{32}P]dCTP (50 pmol, 2000 Ci/mmol) (Amersham) to each 25 μl reaction mixture. Then 5 μl of each amplified sample were digested with the combination of restriction enzymes so that the entire coding region was resolved into several overlapping fragments of a size between 100 and 400 bp. After digestion, 5 μl of the digest were withdrawn and mixed with an equal volume of 95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.5% xylene cyanol. The mixture was heat-denatured at 95°C and snap-frozen on dry ice. Four microlitres of each sample per lane before and after heat denaturation were applied to a 6% polyacrylamide gel (32 × 40 × 0.03 cm; 0.5 cm/lane) containing 90 mM Tris–borate buffer (pH 8.3) and 10% glycerol. Electrophoresis was performed at 3–6 W for 10–16 h at 25°C. The gel was dried on filter paper and exposed to Fuji RX 100 X-ray films at −80°C for 12–48 h with an intensifying screen. Whenever abnormal mobility was detected, the corresponding PCR product was further investigated by nucleic acid sequencing.

Nucleotide sequencing

When biotinylated primers were used in the PCR, the nucleotide sequence was determined by direct solid phase sequencing after obtaining single-stranded DNA by denaturation in 0.1 M sodium hydroxide and separation of the strands by the use of streptavidin-coated magnetic beads (Dynabeads® M-280) and a magnetic particle concentrator (Dyndal MPC® -E) according to the instructions of the distributor. Sequencing was performed by dideoxy chain termination using T7 DNA polymerase (Sequenase™, Biochemical Corp.). When direct sequencing was not informative, the PCR fragments were cloned into phage M13 and sequenced from single-stranded M13 DNA. In all cases exon 2 was sequenced from genomic DNA. Other portions corresponding to abnormal fragments detected by SSCP analysis were sequenced from cDNA.

Analysis of the genomic PIG-A gene

A total of 6×10^5 plaques of a human genomic library in λ2001 (prepared from the Jurkat T-cell line) were screened with the same probe used in the Northern analysis. Two different phages hybridizing to the PIG-A probe were isolated and mapped by restriction enzyme digestion. A 1.2 kb *Hin*dIII fragment and a 2.3 kb *Hin*dIII–*Eco*RI fragment that contained nucleotides −62 to 715 of the cDNA without interruption by any intron sequences were subcloned into pUC12 and 13, and partially sequenced in order to obtain the intron–exon boundaries of exon 2. Exon 1, containing the first 23 nucleotides of the 5′ end of the published cDNA sequence, was located in a 3.2 kb *Hin*dIII fragment 5′ to exon 2.

Acknowledgements

We thank C.Corcoran, T.Vulliamy, J.V.Melo and N.Cross for helpful advice, W.Watkins and N.Dibb for discussions, and D.Stevens for technical assistance. M.B. is supported by grants from the Walter Honegger Stiftung (Zurich) and the Overseas Research Students Awards Scheme (UK). P.H.

is supported by the Wellcome Trust. This work has also received funding from a grant from the Special Health Authority and the Medical Research Council.

References

Armstrong,C., Schubert,J., Ueda,E., Knez,J.J., Gelperin,D., Hirose,S., Silber,R., Hollan,S., Schmidt,R.E. and Medof,M.E. (1992) *J. Biol. Chem.*, **267**, 25347–25351.

Baserga,S.J. and Benz,E.J. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 2056–2060.

Bessler,M., Hillmen,P. and Luzzatto,L. (1992) *Blood*, **80**, 844–845.

Bingham,P.M., Chou,T.-B., Mims,I. and Zachar,Z. (1988) *Trends Genet.*, **4**, 134–138.

Chirgwin,J.M., Przybyla,A.E., Macdonalds,R.J. and Reutter,W. (1979) *Biochemistry*, **18**, 5294.

Dacie,J.V. and Lewis,S.M. (1972) *Ser. Haemat.*, **3**, 3–23.

De Sandre,G., Ghiotto,G. and Mastella,G. (1956) *Acta Med. Patavinia*, **16**, 310–335.

Gonzalez,I.L., Gorski,J.L., Campen,T.J., Dorney,D.J., Erickson,J.M., Sylvester,J.E. and Schmickel,R.D. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 7666–7670.

Green,M.R. (1991) *Annu. Rev. Cell Biol.*, **7**, 559–599.

Hansch,G.M., Schonermark,S. and Roelcke,D. (1987) *J. Clin. Invest.*, **80**, 7–12.

Hillmen,P., Bessler,M., Bungey,J. and Luzzatto,L. (1993a) *Som. Cell Mol. Genet.*, **19**, 123–129.

Hillmen,P., Bessler,M., Crawford,D. and Luzzatto,L. (1993b) *Blood*, **81**, 193–199.

Hillmen,P., Bessler,M., Mason,P., Watkins,W. and Luzzatto,L. (1993c) *Proc. Natl Acad. Sci. USA*, **90**, 5272–5276.

Holander,N., Solvaray,P. and Springer,T.A. (1988) *J. Immunol.*, **141**, 4238–4290.

Margolskee,R.F., Kavathas,P. and Berg,P. (1988) *Mol. Cell. Biol.*, **8**, 2837–2847.

Mason,P.J., Enver,T., Wilkinson,D. and Williams,J.G. (1993) In Hames,B.D. and Higgins,S.J. (eds), *Gene Transcription*. Oxford University Press, Oxford, pp. 5–65.

McCallum,F.S., Eduard,B. and Maden,H. (1985) *Biochem. J.*, **232**, 725–733.

Miyata,T., Takeda,J., Iida,J., Yamada,N., Inoue,N., Takahashi,M., Maeda,K., Kitani,T. and Kinoshita,T. (1993) *Science*, **259**, 1318–1320.

Nicholson-Weller,A., March,J.P., Rosenfeld,S.I. and Austen,K.F. (1983) *Proc. Natl Acad. Sci. USA*, **80**, 5066–5070.

Oni,S.B., Osunkoya,B.O. and Luzzatto,L. (1970) *Blood*, **36**, 145–152.

Oshima,Y. and Gotoh,Y. (1987) *J. Mol. Biol.*, **195**, 247–259.

Pangburn,M.K., Schreiber,D. and Muller-Eberhard,H.J. (1983) *Proc. Natl Acad. Sci. USA*, **80**, 5430–5434.

Persico,M.G., Viglietto,G., Martini,G., Toniolo,D., Paonessa,G., Moscatelli,C., Dono,R., Vulliamy,T. and Luzzatto,L. (1986) *Nucleic Acids Res.*, **14**, 2511–2521 and 7822.

Rosse,W.F. (1990) *Blood*, **75**, 1595–1601.

Rotoli,B. and Luzzatto,L. (1989) *Semin Haematol.*, **26**, 201–207.

Rotoli,B., Robledo,R., Scarpato,N. and Luzzatto,L. (1984) *Blood*, **64**, 847–851.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory, NY.

Selvaraj,P., Rosse,W.F., Silber,R. and Springer,T.A. (1988) *Nature*, **333**, 565–567.

Smith,C.W.J., Nadal-Ginard,B. and Patton,J.G. (1989) *Annu. Rev. Genet.*, **23**, 527–577.

Takahashi,M., Takeda,J., Hirose,S., Hyman,R., Inoue,N., Miyata,T., Ueda,E., Kitani,T., Medof,M.E. and Kinoshita,T. (1993) *J. Exp. Med.*, **177**, 517–521.

Takeda,J., Miyata,T., Kawagoe,K., Iida,Y., Endo,Y., Fujita,T., Takahahi,M., Kitani,T. and Kinoshita,T. (1993) *Cell*, **73**, 703–711.

Urlaub,G., Mitchell,P.J., Ciudad,C.J. and Chasin,L.A. (1989) *Mol. Cell. Biol.*, **9**, 2868–2880.

Received on September 3, 1993; revised on October 8, 1993