Alternative Splicing of a Novel Glycophorin Allele GPe(GL) Generates Two Protein Isoforms in the Human Erythrocyte Membrane

By Cheng-Han Huang, Olga O. Blumenfeld, Marion E. Reid, Ying Chen, Geoff L. Daniels, and Elizabeth Smart

The Henshaw antigen (synonym: He or MNS6) is carried by an altered form of glycophorin B (GPB), but the molecular basis for its variable expression or quantitative polymorphism remains largely undefined. We report here the identification and analysis of a novel glycophorin He allele, GPe(GL), which gives rise to the expression of two protein isoforms in the erythrocyte membrane. In addition to the nucleotide changes defining the epistopic sequence of He, a single C-to-G nucleotide transversion in exon V coding for the membrane domain was found to cause aberrant RNA splicing by creating a new acceptor splice site. In addition, for the phenotypic diversity of membrane glycophorins, a T-to-G transversion at the membrane domain was found to cause aberrant RNA two GPHe isoforms and thus delineate a new mechanism for the phenotypic diversity of membrane glycophorins.

Materials and Methods

Blood sample and hemagglutination testing. The He⁺ blood sample was obtained from a native South African (GL) who was found by screening with rabbit anti-He. Monoclonal anti-He (22G4) and anti-N (14E) were supplied by Gamma Biologicals, Inc (Houston, TX), and anti-S-like (monoclonal antibody [MoAb] 148) was supplied by Immucor (Norcross, GA). Hemagglutination was performed using standard serological testing in test tubes.

Immunoblotting analysis of membrane glycophorins. Immunoblotting was performed as previously described, and peroxidase-conjugated antimouse Ig was used as the secondary antibody.

Southern blot analysis of glycophorin genes. Genomic DNA was isolated from cells in the buffy coat of GL’s blood sample. Both He⁻ and He⁺ DNAs were used as controls. Southern blot analysis following restriction endonuclease digestion was performed as described.

Transcript analysis by reverse-transcriptase polymerase chain reaction (RT-PCR) and cDNA sequencing. Total RNA was isolated from reticulocytes hemolysates and extracted with Trizol reagent.

From Lindsley F. Kimball Research Institute, New York Blood Center; the Department of Biochemistry, Albert Einstein College of Medicine, Bronx, NY; Blood Group Unit, Bristol, UK; Natal Blood Transfusion Service, Durban, South Africa.

Submitted October 29, 1996; accepted February 12, 1997.

Supported by Grants No. HL54459 and GM16389 from the National Institutes of Health, Bethesda, MD and NBF94-10 from the National Blood Foundation (Bethesda, MD).

Address reprint request to Cheng-Han Huang, MD, PhD, Lindsley F. Kimball Research Institute, New York Blood Center, 310 E 67th St, New York, NY 10021.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

© 1997 by The American Society of Hematology.

0006-4971/97/9001-0029$3.00/0
using the supplier’s protocol (BRL, Gaithersburg, MD). RT-PCR of
glycophorin mRNAs was performed as previously described.12 Four
primers were used for synthesis and amplification of GPB and GPHe
cDNAs: P1, N-specific, 5′-GAAATTTGACATCATCAGCTT-3′ (sense, exon II); P2, He-specific, 5′-AAATGGACATAT-
CATCAGCTG-3′ (sense, exon II); P3, GPB-specific, 5′-GTTCATG
GAAGATCAGGCCAGCAT-3′ (antisense, exon V); and P4, GPB-
specific, 5′-GAAGATGAAATGGAAGGAGCATG-3′ (antisense, 3′-untranslated region or 3′-UTR). Two primers were used for syn-
thesis and amplification of GPA cDNA: P5, 5′-GGAATTCCAGCT-
CATGATCAGCTGAG-3′ (sense strand attached with an EcoRI
site GAATTC, exon I) and P6, 5′-TCCACATGGTTGTTGGAAGCAGAT-
TCC-3′ (antisense, exon VII or 3′-UTR).

For the analysis of GPB and GPHe transcripts, total RNAs were con-
verted into single-stranded cDNAs using 20 ng of P4 in 10 µL
of reaction volume (42°C for 1 hour). An aliquot (2.5 µL) of the
cDNA products was then amplified using either P1+P3 or P2+P3
combination. This approach avoids coamplification of GPB and
GPHe, thus producing specific DNA templates suitable for direct
sequencing. PCR was run for 30 cycles as follows: 94°C for 60
seconds, 55°C for 45 seconds, and 72°C for 30 seconds; the last step
was 55°C for 2 minutes and 72°C for 7 minutes. The resultant
cDNA products were analyzed by agarose gel electrophoresis and visualized
by ethidium bromide staining. After purification by polyacrylamide
gel electrophoresis and elution in TE buffer (10 mmol/L Tris-HCl,
0.1 mmol/L EDTA, pH 8.0), cDNA templates were recovered and acid sequence.
directly cycle-sequenced on an automated DNA sequencer using
chain terminators tagged with fluorescent dyes (Applied Biosystem,
Foster City, CA).

Amplification of genomic sequences by allele-specific PCR.
To confirm the results of cDNA sequence analysis and determine the exon-intron structures, allele-specific PCR16 of GPHe(GL)
was performed using its 9.2- and 4.3-kb
Msp
* I fragments as DNA
templates. Four fragments, with each encompassing one exon,
were amplified by the following combinations of primer pairs: for exon II, 5′-GAAGATATATATAAAAAGCGCTTACG-3′
(sense, intron 1) and 5′-GAACCTGAGTATTATTTCTGTGA-3′
(antisense, intron 2); for exon III, 5′-CATCTCTTGGTCTTC-
TTTTCACTTC-3′ (sense, intron 2) and 5′-AGAAGCTTG
CATGATCAGCTGATG-3′ (antisense, intron 3); for exon IV, 5′-CCTTACATGATGACATATTCAATG-3′ (sense, intron 3) and 5′-GAAATG
AGATGATGTTAATAATGGGAGCAAGCT-3′ (antisense,
intron 4); and for exon V, 5′-TGGTCAATTATTACGTTTTCACT-
TAGT-3′ (sense, intron 4) and 5′-CTGTATTTCTTTGGATAG
TAACTGT-3′ (antisense, intron 5). The amplified genomic prod-
ucts were purified and then sequenced, as mentioned above. The exon-intron junction sequences were assessed to derive the score
for the various splice sites, as described.17

RESULTS
Hemagglutination. RBCs of proband GL typed as M-
N+S+u+U+ and had a strong expression of the He antigen
as determined by reactions with mouse monoclonal, rabbit
polyclonal, and human polyclonal anti-He antibodies. Other
antigen typings were unremarkable. The direct antiglobulin
test was negative. Unexpectedly, after α-chymotrypsin treat-
ment, the proband’s RBCs still reacted strongly with mono-
clonal anti-He, even though the S antigen was no longer
detectable on these cells.

Detection of two membrane He+ protein isoforms by im-
umnoblotting. Immunoblots with a monoclonal anti-He
showed an expression in the GL’s RBC membranes of two
He+ glycophorin forms (designated GPHe-1 and GPHe-2).
GPHe-1 was equivalent to GPB in molecular size (Mr
25,000), whereas GPHe-2 was a new fast-moving band with an
Mr approximately 2,700 daltons less than GPB (Fig 1A, lane 3). The
latter glycoprotein species was absent in RBCs from both He− and He+ controls (Fig 1A, lanes 1 and 2). Immunoblotting with an anti-N MoAb detected GPB in the
normal control and GL (Fig 1B), but the GPHe-2 species was
nonreactive with the antibody (Fig 1B, lane 2). The
GPHe-2 band was also detectable by anti-M′ but not by anti-
N (nonreactive with M−, He+ RBCs), anti-s or MoAb148 that is anti-S− like (data not shown). Together, these results
indicated that the proband is heterozygous for a normal GPB
gene and an altered GPHe gene producing two He+ mem-
brane proteins.

To further explore the structural differences between the
two He+ protein isoforms, α-chymotrypsin digestion of
RBCs followed by immunoblotting analysis with anti-He
was performed. The results showed that GPHe-1 from both
He+ control and GL was α-chymotrypsin−cleavable (Fig
1C, lanes 7 and 9), whereas GPHe-2 from GL was not (Fig
1C, lane 9). This finding indicated that the enzyme cleavage
sites in the latter were either absent or conformationally
perturbed due to a truncation of the corresponding amino
acid sequence.

Genetic polymorphism of He detected by Southern blot-
ing. In proband GL, Msp I blots hybridized with glycopho-
rin cDNA and genomic DNA probes showed a gross alter-
ation in the GPB gene but not the GPA gene (Fig 2,
autoradiograms at left). In addition to a 13.5-kb band for the
normal GPB gene, there were two new bands of 9.2-kb and
4.3-kb in size that originated from the GPB gene on the
homologous chromosome. These two bands were previously
shown to arise by the introduction of a unique Msp I cleavage
site into the second intron of GPB gene, perhaps through
a gene conversion-like event.12 Analysis of unrelated He−
and He+ genomic DNAs showed that only the He+ samples
display this polymorphism (gels not shown), suggesting its
tight linkage with He. A genetic marker, the 7.2-kb Msp I
band, known to be tightly associated with S-s−,14 was also
seen in GL. These data confirmed that the diploid genome
of GL carries a normal GPB gene and an altered GPHe gene
designated GPHe(GL)), Fig 2, right).

Splicing pattern of GPHet(GL) transcripts. To establish
whether the two GPHe protein isoforms are derived from
the same gene or two different genes, RT-PCR and cDNA
sequencing were used to investigate the composition and
structure of glycophorin transcripts (Fig 3A). Compared to
normal controls, no abnormality was found in GL with re-
gard to GPA (data not shown) and GPB transcripts (Fig 3B,
lanes 1 and 2). However, GPHe(GL) cDNAs amplified with
P2 and P3 consisted of four species, one comparable to,
while others smaller than, GPB in size (Fig 3B, lane 4). Nucleotide
sequencing identified the pattern of various exon-

From www.bloodjournal.org by guest on October 27, 2017. For personal use only.
Fig 1. Immunoblotting analysis of RBC membrane glycophorins. (A) Immunoblots probed with monoclonal anti-He. Lane designations are: 1, He−, He negative control; 2, He+, He positive control; and 3, proband GL. He monomer is seen in He+ but not He−, whereas two monomeric components of GPHe are present in GL (indicated by He and an arrow). Note that the upper bands seen in lane GL are due to overload and formation of homodimers and heterodimers with other glycophorins. (B) Immunoblots probed with monoclonal anti-N. Note that RBC membranes from proband GL contain a GPB form comparable to that from He− control (B denotes the monomeric form). (C) RBCs were treated with α-chymotrypsin and then membranes were prepared for immunoblotting with anti-He. A minus sign (−): untreated, and a plus sign (+): chymotrypsin-treated. This enzyme digestion almost completely removed the GPHe monomer with a size comparable to GPB but not at all the lower-molecular-weight GPHe species (arrow-indicated).

Fig 2. Southern blot analysis of glycophorin genes. Left, genomic DNAs were digested with Msp I and hybridized with GPA cDNA and intron 2 and 3 probes, as indicated. Lanes 1, He− control (M+N+He+S–s–U–); 2, He+ control (M+N+He+S–s–U–); and 3, proband GL (M+N+He+S+s+U+). The size (kb) and gene origin of various bands are indicated at left and right margins, respectively. Arrow points to the 7.2-kb marker band known to be tightly linked with S–s. Right, the probable genotypes for the GYP A locus in the three subjects are schematically shown in which the GPE gene is omitted for the sake of simplicity. He(S–U–) denotes nondeleted haplotype and S–s–U– GPB-deleted haplotype. GL is a heterozygote for the He(S–U–) haplotype. The GPHe gene-specific 9.2- and 4.3-kb bands originating from the introduction of a unique Msp I site into intron 2 is indicated at bottom.
Fig 3. RT-PCR analysis and splicing pattern of GPHe(GL) transcripts. (A) Strategy for gene-specific RT-PCR analysis. The correspondence of exons to different protein domains and the location of primers for cDNA synthesis and amplification are shown. (B) Agarose gel electrophoresis of GPB and GPHe(GL) cDNAs. M is the HaeIII-cleaved size marker of φX174 DNA. Lanes 1 and 3, He control; and lanes 2 and 4, GL. When primers P2 and P3 were used, multiple cDNA forms are seen in GL, while no product is found in control. Lanes 5 through 8 show reanalysis of the PAGE-puriﬁed cDNAs of GPHe-2, 1, 3, and 4 observed in lane 4. (C) The pattern of exon-exon connections in GPHe-1 to 4 isoforms. Like GPB, exon III of GPHe(GL) is also a pseudoexon (denoted c) attached with a defective donor site, TT. GPHe-1 and GPHe-2 are products of in-frame splicing, whereas GPHe-3 and GPHe-4 are products of out-of-frame splicing with the same frameshift and premature termination (see Fig 4A). The newly created acceptor site (ACAG) and termination codon TGA in exon V are indicated. The residual exon V sequence spliced into GPHe-3 and GPHe-4 is blackened.

In GPHe-3, exon IV was retained, but it was joined to the 3′ end of exon V, thereby resulting in a premature stop codon. GPHe-4 lacked both exon IV and a large portion of exon V, causing a further truncation of the coding sequence but a same premature termination as seen in GPHe-3 (see below).

Nucleotide and deduced amino acid sequences of GPHe isoforms. Figure 4 shows the nucleotide and deduced amino acid sequences of GPB and the four GPHe transcripts obtained from GL’s erythroid cells. The GPB sequence should encode the N and s antigens since it contains Leu, Glu, and Thr at positions 1, 5, and 29, respectively. All GPHe(GL) forms were predicted to carry Trp, Ser, and Gly residues at positions 1, 4, and 5, which define the He antigen. It is also apparent from comparison with GPB that all GPHe transcripts carry the same two silent nucleotide substitutions in exon II, nt 39T-to-C, and nt 45A-to-G (Fig 4A). These results confirmed that all GPHe(GL) cDNAs originated from the same precursor messenger RNA (Pre-mRNA) by alternative splicing (Fig 3C).

Comparison of GPB and GPHe-1 further showed that there was a single C-to-G base change in exon V that resulted in a Thr-to-Ser substitution at position 65 (Fig 4B). In conjunction with the upstream pyrimidine-rich sequence, this guanyl nucleotide created a new cryptic acceptor splice site, ACAG (Fig 4B, overlined and boxed). As shown, the sequence of exon IV was completely excluded from both GPHe-2 and 4, whereas the sequence of exon V was partially excluded from GPHe-3 and 4, exactly at the position of AG dinucleotide. These results showed a partial activation of the new acceptor splice site and partial inactivation of the normal splice sites.

While GPHe-1 should contain 72 amino acids with a Met residue at position 29 for the S antigen, the truncated GPHe forms lack either the Met residue or the membrane domain or both because of the different skipping events involving exon IV and/or exon V (Fig 4A). GPHe-2 maintained a correct open reading frame and thus was predicted to encode a polypeptide with 59 amino acids; however, exclusion of all GPHe transcripts carry the same two silent nucleotide substitutions in exon II, nt 39T-to-C, and nt 45A-to-G (Fig 4A). These results confirmed that all GPHe(GL) cDNAs originated from the same precursor messenger RNA (Pre-mRNA) by alternative splicing (Fig 3C).
IV was also identified. Although it would only reduce the splice site score from 90 to 84, the aberrant splicing events associated with GPHe(GL) pre-mRNA could be the combined effects of this mutation and the creation of a new acceptor site within exon V. Figure 5B shows the exon-intron junction sequences of the GPHe (GL) gene and comparison of scores for the new and original splice sites. This assessment indicates that the new acceptor site has a score comparable to the donor site in intron 4 (84 vs 89), but much higher than that of the adjacent acceptor site that is only 76-base pairs apart (84 vs 64).

**DISCUSSION**

In this study, we describe the identification and molecular analysis of a novel glycophorin allele, GPHe(GL), which produces two He+ protein isoforms in the erythrocyte mem-

---

**Fig 4.** Nucleotide and deduced amino acid sequences of the four GPHe(GL) isoforms. (A) Nucleotide sequences of the four GPHe(GL) cDNA species. The GPB sequence from GL is listed for comparison. Nucleotide differences are marked by stars and deleted sequences in the cDNA products by dashes. The new acceptor splice site ACAG is boxed and its upstream pyrimidine-rich sequence indicated by a straight line. Exons are numbered and their boundaries denoted by triangles. Termination codons (TGA) are underlined. (B) Alignment of the deduced amino acid sequences for GPHe(GL) isoforms. Amino acid variations are denoted by stars and deleted sequences by dashes. Gaps interrupting the amino acid sequence pertain to the exon-exon junctions. The total amino acid number of each putative polypeptide is indicated. Note that GPHe-3 and 4 are premature chain termination products because of a shift in the open reading frame.

---

**Fig 5.** Genomic structure of the GPHe(GL) allele and nucleotide sequences of exon-intron junctions. (A) Strategy for genomic amplification of exons and their flanking sequences. Four segments with each covering a unique exon were amplified from either the 9.2- or 4.3-kb Msp I fragment. Segment e, which overlaps a and b, was amplified from total genomic DNA using the He-specific primer P2 and the reverse primer for the b segment (see Materials and Methods). As shown, the GPHe gene from GL also is a GPB-A-B hybrid in the exon II-intron 2 junction region. 12 (B) Exon-intron junction structures and splice site scores for the determined acceptor and donor splice sites (overlined). Exon sequences are in capital letters and their encoded amino acid residues are shown. Dots denote omission. Two point mutations, a g$\rightarrow$rt transversion in the 5* donor element of exon III and a t$\rightarrow$rg transversion at position 6 near the exon IV acceptor site, are indicated by arrows. The newly created acceptor splice site in exon V and its score are also shown. Exon sequences are in capital letters and their encoded amino acid residues are shown. Dots denote omission. Two point mutations, a g→rt transversion in the 5' donor site attached to exon III of the GPHe(GL) gene (Fig 3C, TT instead of GT) and of 9 nucleotide differences identified by cDNA sequencing (Fig 4A). Significantly, a point mutation (t$\rightarrow$g) in the pyrimidine stretch of the acceptor site of exon IV was reversed. Although it would only reduce the splice site score from 90 to 84, the aberrant splicing events associated with GPHe(GL) pre-mRNA could be the combined effects of this mutation and the creation of a new acceptor site within exon V. Figure 5B shows the exon-intron junction sequences of the GPHe (GL) gene and comparison of scores for the new and original splice sites. This assessment indicates that the new acceptor site has a score comparable to the donor site in intron 4 (84 vs 89), but much higher than that of the adjacent acceptor site that is only 76-base pairs apart (84 vs 64).
brane. We have shown that this gene codes for He, S, and U antigens but contains two nucleotide transversions. The T-to-G mutation is located at −6 position of the acceptor splice site for exon IV, whereas the C-to-G substitution resides in exon V encoding the membrane-spanning segment. Evidence is presented that the latter point mutation creates a new acceptor splice site whose partial activation causes aberrant RNA splicings. We have also shown that the probe carries a normal GBP gene on the homologous chromosome that encodes N, s, and U antigens. These results correlate alternative splicing with the expression of two GPHe isoforms and delineate a new mechanism underlying the polymorphism of He antigen.

The coexistence of both He and S coding sequences in the GPHe-1 transcript conforms to the expression of corresponding antigens on the variant erythrocytes and to the occurrence of a He+ species that comigrated with normal GBP on immunoblots. GPHe-2, rather than GPHe-3 or GPHe-4, is more likely to represent the lower Mr He+ protein species expressed in the RBC membrane, because GPHe-2 is a major transcript and encodes a polypeptide with an intact transmembrane segment. In GPHe-2, the absence of an exon IV sequence also is consistent with the lack of reactivity of its protein with anti-S–like MoAb. Moreover, in view of the potential cleavage sites of α-chymotrypsin in GBP,19 such a deletion readily explains the abolished enzyme digestion of the GPHe-2 protein. With regard to GPHe-3 and GPHe-4, the low level of transcript expression, the deletion of a large internal sequence, and the truncation of a major portion of the membrane domain would impede the disposition of their putative protein products in the plasma membrane. Detailed transfection studies using the GPHe(GL) cDNAs are required to provide a definitive proof for the gene-phenotype correlation.

With the GPHe gene analyzed here, three different forms of the He antigen have now been characterized at the molecular level. In coexpression of He and St7, genetic recombination via the third introns of GPHe and GPA genes leads to a hybrid gene GPHe(St*) encoding both antigens.12 As to the linkage of He with S-s-U*, the responsible GPHe(P*) gene contains two splicing mutations, one being identical with that described here and the other being a G-to-T change at +5 position of the consensus donor splice site.12 In a concerted fashion, the two mutations cause a complete skipping of exon V, resulting in a frameshift and an elongated new hydrophobic sequence for membrane anchoring. Interestingly, the protein product only displays the He reactivity, although it contains a linear sequence for both He and S (and perhaps U) antigens. In contrast, the existing example shows a different splicing pattern in which the joining of exon IV and/or exon V is affected due to the partial activation of the new acceptor splice site and the partial inactivation of the normal splice sites. Accordingly, the exon IV-deleted transcript GPHe-2 lacks the coding sequence for the S and U antigens and thus expresses only the He antigen. These findings support the proposal that diverse mechanisms may underlie the phenotypic variation of He, whether qualitatively or quantitatively.10

Pre-mRNA splicing is a hierarchical cellular process by which splice sites at the exon-intron boundaries can be selected precisely.20,21 The information essential for the selection includes the nearly invariant GT and AG elements at the 5′ and 3′ ends of introns and a branch point close to the 3′ acceptor splice site.17 A survey of human disease genes has shown that point substitutions in GT and AG account for the majority of splicing mutations, constantly causing aberrant RNA splicing.22 However, the effects on pre-mRNA splicing of those mutations not occurring in the predetermined GT-AG elements at exon-intron junctions are not readily predictable. The multiple exon skipping events described here apparently occurred as the result of the combined effects of the T-to-G change close to the exon IV acceptor site and the C-to-G mutation in exon V leading to a new acceptor site. With regard to exon IV and/or exon V partial skipping, it remains to be investigated whether the two altered nucleotides exert their effects concertedly or independently during pre-mRNA splicing. Nevertheless, the selection of the new acceptor site in exon V appears to be related to its sequence context, spatial distribution, and balanced competition with other splice sites. First, the AG element is coincident with the occurrence of a stretch of pyrimidine-rich sequence and thus shows a high similarity to the consensus acceptor site.17 Second, due to a proximity or position effect,23,24 the new splice site might compete with the upstream acceptor and downstream donor sites and thus result in alternative splicing. Third, the selective usage of the new acceptor site could be related to the inevitable emergence of two premature stop codons in exon V, as nonsense mutations are known determinants of splice site selection.27 Along with our further understanding of the molecular basis for He variation, it should be possible to determine how different combinations of mutations affect the processing of GPHe-related primary transcripts by direct experimentation involving splicing assays performed on model pre-mRNA constructs.

ACKNOWLEDGMENT

We thank John Moulds from Gamma Biologicals for rabbit anti-He and monoclonal anti-He and anti-N, and Susan Rolih from Immucor for human anti-He and monoclonal anti-S–like. We are grateful to Dr Colvin Redman for reading the manuscript and helpful discussion. Technical assistance from Rita Batts, Gregory Halverson, and Frances Green is highly appreciated. Thanks are also due to Tellerv Huima-Byron and Yelena Oskov for making photoprints.

REFERENCES


7. Lomas C, Daniels GL, Green CA, Kelly L, Moulds M: The first He/St(a+) donor: Immunoblotting determined the presence of He and St+ on one novel glycophorin molecule. Transfusion 32:54S, 1992 (abstr)


11. Huang C-H, Lomas C, Daniels G, Blumenfeld OO: Glycophorin He(St') of the red blood cell membrane is encoded by a complex hybrid gene resulting from two recombinational events. Blood 83:3369, 1994


Alternative Splicing of a Novel Glycophorin Allele GPHe( GL) Generates Two Protein Isoforms in the Human Erythrocyte Membrane

Cheng-Han Huang, Olga O. Blumenfeld, Marion E. Reid, Ying Chen, Geoff L. Daniels and Elizabeth Smart

Updated information and services can be found at:
http://www.bloodjournal.org/content/90/1/391.full.html

Articles on similar topics can be found in the following Blood collections

Information about reproducing this article in parts or in its entirety may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at:
http://www.bloodjournal.org/site/subscriptions/index.xhtml