**A Large Deletion Within the Protein 4.1 Gene Associated With a Stable Truncated mRNA and an Unaltered Tissue-Specific Alternative Splicing**

By N. Dalla Venezia, P. Maillet, L. Morlé, L. Roda, J. Delaunay, and F. Baklouti

Protein 4.1 is a major protein of the red blood cell (RBC) skeleton that laminates the inner surface of the plasma membrane (for review, see Lux and Palek, Delaunay, and Conboy). It contributes to the skeleton anchoring to the membrane through interactions of its 30-kD N-terminal domain and to the spectrin-actin lattice through its 10-kD domain. We describe here the molecular basis of a heterozygous hereditary elliptocytosis (HE) associated with protein 4.1 partial deficiency. The responsible allele displayed a greater than 70-kb genomic deletion, beginning within intron 1 and ending within a 1.3-kb region upstream from exon 13. This deletion encompassed both erythroid and nonerythroid translation initiation sites. It accounts for the largest deletion known in genes encoding proteins of the red blood cell membrane. The corresponding mRNA was shortened by 1727 bases, due to the absence of exons 2 to 12. Nevertheless, this mRNA was stable. It showed a similar pattern in lymphoblastoid cells as in reticulocytes. Differential splicing of exons within the undeleted region remained regulated in a tissue-specific manner. Exons 14, 15, and 17a were absent from both reticulocyte and lymphocyte mRNAs, whereas exon 16 was present in reticulocytes but absent from lymphocytes. Thus, differential splicing on a local scale was not dependent on the overall structure of protein 4.1 mRNA in this particular instance.

© 1998 by The American Society of Hematology.
mRNA studies. Total reticulocyte RNA was obtained from peripheral blood as previously described.24 Lymphoblastoid cell lines were kindly established by Prof G. Lenoir (International Agency for Research on Cancer, Lyon, France) from blood sample of proband I.2. Total RNA was extracted from cultured cells according to Chomczynski and Sacchi.26

Reverse transcription (RT) and PCR were performed essentially as previously referred to22 using sequence specific primers (Table 1). The 30 to 35 cycles of amplification were performed as follows: denaturation at 94°C for 45 seconds, annealing at 56°C for 30 seconds or at 60°C for 45 seconds, extension at 72°C for 60 seconds.5 8RACE was performed to characterize the shortened protein 4.1 mRNA revealed by Northern blot analyses.24 Either random hexamers or primer XI in RT experiments and two pairs of primers (I/XII) or (I/X) in PCR experiments (Table 1 and Fig 1) were used. The reactions yielded a major 522-bp fragment (not shown) and a 418-bp fragment, respectively, corresponding to the abnormal 4.1 allele (Fig 2A). Because the size of the amplified fragment would exceed 1 kb (maximum size recommended by the manufacturer) in the normal allele, no PCR products were expected from these amplifications in control mRNA. Indeed, no PCR products were obtained from the control (Fig 2A). 5’RACE using an upstream reverse primer within exon 13 (primer IX, Fig 1 and 4362 DALLA VENEZIA ET AL

RESULTS

A shortened protein 4.1 mRNA. 5’RACE was performed to characterize the shortened protein 4.1 mRNA revealed by Northern blot analyses.24 Either random hexamers or primer XI in RT experiments and two pairs of primers (I/XII) or (I/X) in PCR experiments (Table 1 and Fig 1) were used. The reactions yielded a major 522-bp fragment (not shown) and a 418-bp fragment, respectively, corresponding to the abnormal 4.1 allele (Fig 2A). Because the size of the amplified fragment would exceed 1 kb (maximum size recommended by the manufacturer) in the normal allele, no PCR products were expected from these amplifications in control mRNA. Indeed, no PCR products were obtained from the control (Fig 2A). 5’RACE using an upstream reverse primer within exon 13 (primer IX, Fig 1 and

Table 1. Oligonucleotides Used in RT or PCR Experiments

<table>
<thead>
<tr>
<th>Primer Sequence</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. GC (GAATTCC) CTCTGAAAGGTTC-CAGAATCGATAG</td>
<td>AS Anchor*</td>
</tr>
<tr>
<td>II. AGAAGTGGTTTGGTCAGGGGTC</td>
<td>S</td>
</tr>
<tr>
<td>III. GGACCTCTACCGCTGACC</td>
<td>AS</td>
</tr>
<tr>
<td>IV. CCGAAGCGGCCGGCCTCCGC</td>
<td>S 61-79</td>
</tr>
<tr>
<td>V. TTCTCTGGTCCTCTGAGC</td>
<td>AS 680-660</td>
</tr>
<tr>
<td>VI. AGCAGTATGAAAGTACCATCGG</td>
<td>S 1558-1579</td>
</tr>
<tr>
<td>VII. GGTACCTGCTGAGCCTCG</td>
<td>AS 1800-1781</td>
</tr>
<tr>
<td>VIII. CTGCAGATTGCAGCCAGCCA</td>
<td>S 1831-1850</td>
</tr>
<tr>
<td>IX. TCTGTCCTTTAGGGACCACTG</td>
<td>AS 1953-1933</td>
</tr>
<tr>
<td>X. GC (GAATTC) GATCTCCTCTT-GACCTTTG</td>
<td>AS 2180-2162</td>
</tr>
<tr>
<td>XI. GCTCAGCTATGCTGGACATG</td>
<td>AS 2208-2187</td>
</tr>
<tr>
<td>XII. GC (GAATTC) GGTGAGTGGATGT-GATAAGGG</td>
<td>AS 2284-2264</td>
</tr>
<tr>
<td>XIII. AGAAGCCACAGACATGGA</td>
<td>S 2012-2031</td>
</tr>
<tr>
<td>XIV. TCACCTTTCAGACATTGGCATATTCT</td>
<td>AS 2392-2366</td>
</tr>
<tr>
<td>XV. GACAGAAGTCGAAAGGAGATC</td>
<td>S 2160-2180</td>
</tr>
<tr>
<td>XVI. AAGGGCAGATGGTGGGAGTGG</td>
<td>AS 2774-2755</td>
</tr>
</tbody>
</table>

The location of nucleotides is given within the sequence published by Conboy et al.25
Abbreviations: S, sense strand; AS, antisense strand.
*Sequence complementary to the anchor used in the Amplifinder kit (Clontech).
†Genomic sequences isolated from the W27 clone23; they are located upstream and downstream from exon 1, respectively.23
Table 1) showed a similar pattern, with a major 244-bp band in the proposita (Fig 2A). The amplified product suggested the presence of the 3' region of exon 13, encompassing primer IX sequence, on the abnormal mRNA.

The 418-bp and 244-bp fragments were cloned and 8 positive clones (4 of each size) were sequenced. They all showed an abnormal sequence with exon 13 3' end joined to exon 13 5' end, suggesting a large deletion of about 1727 bp (Fig 2B). This deletion encompasses 10 coding exons (exons 2 through 12, exon 3 being a non-coding sequence23) and therefore removes both upstream and downstream translation initiation codons (Fig 1). To confirm the 5' RACE results, and knowing that the deletion spares exon 1 sequences, we performed a standard RT-PCR experiment using primers IV and IX, designed within exon 1 and exon 13, respectively. Again, a specific PCR product was obtained from the patient's RNA only. Nucleotide sequencing of the 167-bp corresponding band confirmed the large deletion observed in 5' RACE experiments (not shown).

In addition to the large deletion, nucleotide sequencing of the 5' RACE clones showed two discrete changes at the new exon 1/exon 13 boundary (Fig 2B). None of the 8 clones actually showed an intact junction with three repeated CAG codons. Instead, 5 clones showed two of these repeats, and the three remaining clones contained only one CAG codon. Exon 1/intron 1 boundary sequence does not display an alternative donor splicing site with a consensus context. Therefore, it most likely participates to the splicing reaction with the wildtype splicing site. However, the succession of three AG dinucleotides at the intron 13/exon 13 junction offers three potential acceptor splicing sites. It is therefore plausible that the skipping of CAG codons must derive from the use of internal cryptic acceptor sites within exon 13 (Fig 1B).

Deleted protein 4.1 mRNA in lymphoblastoid cells. Lymphoblastoid cell mRNAs originating from the proposita and from a control were investigated. Northern blot analysis using pLym 5' showed a single 6.4-kb band in both the control and the patient. As in reticulocyte 4.1 mRNA, pLym 3' showed a second major 4.4-kb band, in the proposita (Fig 3). Thus, the mutant mRNA in lymphoblastoid cells was also stable and presented a similar 2-kb deletion.

RT-PCR experiments using primers IV and IX showed a 167-bp fragment similar to that found in reticulocyte mRNA; the Taq I digestion provided a rapid mean to probe the nature of this fragment. These data suggest that lymphoblastoid cells also
presented the same 1727-bp mRNA deletion, which connected exon 1 to exon 13.

A large genomic deletion encompassing 11 exons. To test the hypothesis of a gene rearrangement responsible for the mRNA deletion, we first performed a Southern blot analysis of genomic DNA from patients I.2 and II.1 and from the unaffected parent I.1 using a pair of restriction endonucleases Msp I and Xho I. Hybridization with either pLym 3′ or pA2.1 (Fig 1) showed an additional band in both patients occurring at the expense of a normal fragment; pLym 5′ probe did not detect this particular fragment (not shown). These findings suggested that the responsible genomic abnormality would involve the 5′ half of the gene and would end within a genomic region encompassed by pA2.1 sequence.

Genomic DNAs were then further investigated by Southern blot analysis using pA2.1 probe and EcoRI, Xho I, Bgl II, or Msp I restriction enzymes. In 4.1(−) HE patients, several bands presented a diminished signal. On Xho I, Bgl II, and Msp I restriction patterns, additional abnormal-sized bands were obtained (Fig 4A). These experiments, together with the mRNA studies, strongly suggested that the 4.1(−) allele bears a large deletion at its 5′ half.

The 3′ endpoint of the deletion was delineated by Southern blot analysis using shorter cDNA probes: 3EX, EX11/12, and EX13 (Fig 1). The abnormal-sized fragments detected with blot analysis using pA2.1 probe and EX13 probe, which is an exon 13-specific probe, suggesting that the 3′ endpoint must be at the vicinity of this exon. These data, together with our previous genomic clone mapping (Baklouti et al23 and unpublished data), allowed us to locate the 3′ endpoint within a 1-kb intronic region situated 250 bp upstream from exon 13 (Fig 4B).

To define the size and the 5′ endpoint of the genomic deletion, we analyzed the genomic region encompassing exons 1 and 2. Southern blot using EX2 cDNA probe (Fig 1) on genomic DNA digested with Xho I, Bgl II, Apa I, or Spe I (Fig 5) showed half intensity signals in 4.1(−) HE carriers. The comparable signal intensity between the control and the proposita, using pLym3′ cDNA probe, argues against a DNA loading difference and strongly suggests that the reduced signal observed with EX2 probe is related to the gene deletion. These results indicated that the genomic deletion encompassed all the coding exons absent from the mRNA (exons 2 to 12). We further investigated the genomic region surrounding exon 1 using NTER cDNA probe (Fig 1) and fairly rare cutters (BamHI, EcoRV, and Kpn I). The patterns obtained did not show any specific differences between the controls and 4.1(−) HE patients (not shown), suggesting that no additional event such as an inversion or a translocation had occurred during the deleterious event.

We finally attempted to capture the genomic region containing both the 5′ and the 3′ deletion endpoints using a long range PCR and primers IV and IX (Table 1). The PCR experiments, using conditions for up to the 20-kb long fragment, failed to amplify the deleted 4.1 gene. These data, together with the absence of additional bands on Southern blots using either EX2 or NTER probes, as explained above, suggested that the 5′ endpoint must lie within intron 1, several kilobases away from both exon 1 and exon 2. Moreover, the exon-intron organization of protein 4.1 gene study23 presented a gap in the delineation of some intronic sequences: among others, intron 1 appears to be much larger than 25 kb. Therefore, exons 1 and 13 would be separated by at least 70 kb.

In conclusion, 4.1 Annecy allele most likely results from a genomic deletion that spans over 70 kb (Fig 6).

Erythroid and lymphoid expression of alternatively spliced exons. The RT-PCR and 5′RACE experiments described above clearly showed that the chimeric intron, made of the 5′ part of intron 1 and the 3′ part of intron 12, is correctly spliced leading to a junction of exon 1 and exon 13. We next asked whether the large genomic deletion affects the tissue-specific splicing of the remaining exons. Two overlapping RT-PCR experiments were performed on reticulocyte and lymphoblastoid cell RNAs using primers designed within the nondeleted cDNA sequence (Fig 7). These experiments showed that exons 14 and 15, tandemly present in brain,23,27 were absent from normal and shortened mRNAs in both cell types. Exon 17a, recently described as a muscle specific exon,23 was likewise
basis of which is defined at the gene level. The scarcity of protein 4.1 gene mutations elucidated so far contrasts with the profusion of HE generating mutations in SPTA1 and SPTB genes, encoding α- and β-spectrin chains, respectively. This is presumably accounted for by the lack of information at the protein level (protein 4.1 is just manifested by its absence) and the knowledge, only recently acquired, of the protein 4.1 gene structure.

The nucleotide sequence of mutated 4.1 mRNA Annecy is defined by a deletion spanning exon 2 to exon 12 region and removing 1727 bases. Exon 1 and exon 13 are joined, implying that the splicing had normally occurred between these 2 exons. However, 3 nucleotides were missing in the mutated sequence (exon 1)GCCAGCAGCTG(exon 13) as compared with the expected joined sequences (exon 1)GCCAGCAGCTG(exon 13). This stemmed from the use of a cryptic site within exon 13. Such an alternative splicing has already been observed, although as a minor event, in normal mRNA.22 Sequencing analysis showed still another minor species generated, presumably by the skipping of a second CAG triplet (exon 1)GCCAGCAGCTG(exon 13) (Fig 2B), and the use of the remaining AG cryptic site of the repetitive CAG sequence. These data suggest that the sequence of exon 1 is identical to that in the normal 4.1 mRNA sequence and that the 5’ end of exon 13 is submitted to extra splicing events due to repetitive CAG sequences.

Among all 4.1(−) alleles described to date, only those alterations whose expression is confined in the RBCs have been found at the homozygous state; this is the case for allele 4.1 Madrid, mutated in the erythroid translation initiation codon22 and the Algerian 4.1 allele lacking three exons, including the one that carries the downstream codon.29 Use of the upstream AUG codon allows protein 4.1 expression in nonerythroid tissues. However, allele 4.1 Annecy is the only variant lacking both translation initiation sites. One would speculate that such a variant would not exist at the homozygous state.

Large deletions within functional genes usually jeopardize the cytoplasmic mRNA accumulation either because of an altered transcription initiation, defective splicing, RNA transport, or in-frame premature stop codon. We and others have shown that the alteration of the translation initiation site often destabilizes mRNA22,30-32 (for review, see Cooper33). Shortened mRNA 4.1 Annecy, although it potentially encodes for nonfunctional protein(s), remained stable in the cytoplasm. It did so not only in nucleated cells (lymphoblastoid cells), but also in enucleated cells (reticulocytes), long after the nucleus to cytoplasm transport.

We and others have shown that protein 4.1 pre-mRNA is endowed with complex alternative splicing in a tissue and
developmental stage-specific manners. In mRNA 4.1 Annecy, the splicing pattern of exons downstream from the deletion is similar to splicing in normal protein 4.1 mRNA. These findings, consistent with our recent functional studies (Baklouti et al. and unpublished data), suggest that cis-elements involved in exon recognition must lie within targeted exons and their close vicinity intronic sequences.

In this work, we have defined the alteration of allele 4.1 Annecy, which is responsible for 4.1(−) HE. We accurately defined the 1727-bp deletion at the mRNA level. This mRNA was stable and presented the same feature in lymphoblastoid cells. At the gene level, we delineated the first very large genomic deletion encountered in RBC membrane protein genes, extending over 70 kb, starting and ending quite far from the corresponding exons within introns 1 and 12, respectively. Miniallele 4.1 Annecy nevertheless allowed the occurrence of a splicing that connected exons 1 and 13 into a stable and shortened mRNA.

ACKNOWLEDGMENT

The authors thank family PE for their kind cooperation. We thank Prof. G. Lenoir and M. Vuillaume (International Agency for Research on Cancer) for obtaining the lymphoblastoid cell line and Dr. J.P. Magaud for growing the cells.

REFERENCES

21. Tchernia G, Mohandas N, Shohet S: Deficiency of skeletal...
34. Baklouti F, Zhou J, Delaunay J, Huang SC, Benz EJ Jr: Characterization of cis-elements modulating regulated alternative splicing of exons encoding the spectrin/actin binding domain in protein 4.1 pre-mRNA. Blood 84:361a, 1994 (abstr, suppl 1)
A Large Deletion Within the Protein 4.1 Gene Associated With a Stable Truncated mRNA and an Unaltered Tissue-Specific Alternative Splicing

N. Dalla Venezia, P. Maillet, L. Morlé, L. Roda, J. Delaunay and F. Baklouti