Molecular Analysis of Insertion/Deletion Mutations in Protein 4.1 in Elliptocytosis

II. Determination of Molecular Genetic Origins of Rearrangements

John Conboy, Sally Marchesi, Rami Kim, Peter Agre, Y. W. Kan, and Narla Mohandas

Department of Laboratory Medicine and Howard Hughes Medical Institute, University of California, San Francisco, California 94143; Cell and Molecular Biology Division, Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720; Departments of Laboratory Medicine and Pathology, Yale University, New Haven, Connecticut 06510; and Department of Medicine, Johns Hopkins University, Baltimore, Maryland 21205

Abstract

Protein 4.1 is an ~ 80-kD structural protein in the membrane skeleton which underlies and supports the erythrocyte plasma membrane. The preceding companion paper presents a biochemical study of two abnormal protein 4.1 species from individuals with the red blood cell disorder, hereditary elliptocytosis. These variants, "protein $4.1^{68/65}$" and "protein 4.1^{95}," have altered molecular weights due to internal deletions and duplications apparently localized around the spectrin-actin binding domain. Here we use polymerase chain reaction (PCR) techniques to clone and sequence the corresponding mutant reticulocyte mRNAs, and correlate the deletion/duplication end points with exon boundaries of the gene. Protein $4.1^{68/65}$ mRNA lacks sequences encoding the functionally important spectrin-actin binding domain due to a 240 nucleotide (nt) deletion spanning the codons for Lys^{407}-Gly^{486}. Protein 4.1^{95} mRNA encodes a protein with two spectrin-actin binding domains by virtue of a 369 nt duplication of codons for Lys^{407}-Gln^{529}. These deletions and duplications correspond to gene rearrangements involving three exons encoding 21, 59, and 43 amino acids, respectively. The duplicated 21 amino acid exon in the 4.1^{95} gene retains its proper tissue-specific expression pattern, being spliced into reticulocyte 4.1 mRNA and out of lymphocyte 4.1 mRNA. (*J. Clin. Invest.* 1990. 86:524–530.)

Key words: erythrocyte • protein 4.1 • mutations • gene rearrangements PCR

Introduction

Protein 4.1 is a structural protein of 78–80 kD localized to the erythroid membrane skeleton that underlies the red cell plasma membrane. Protein 4.1 plays a critical role in regulating the mechanical properties of the erythrocyte membrane through multiple interactions with integral membrane proteins (band 3 and glycophorin C), with a membrane skeletal protein (spectrin), and with a cytoskeletal protein (actin). Genetic and biochemical evidence (1–3) clearly implicate mutations in protein 4.1 in the etiology of some cases of hereditary elliptocytosis (HE),[1] which vary from mild to severe hemolytic anemia. Chromosome mapping studies have localized the human protein 4.1 gene to a site near the Rh locus that has been closely linked to HE (1p32 → pter, references 3–5). In addition, quantitative deficiencies (1, 2) or qualitative defects (4) in protein 4.1 have been found to co-segregate with the HE phenotype in at least four families, and restriction fragment length polymorphisms were linked to HE and partial 4.1 deficiency in another family (6).

In this report and the preceding companion paper (7), two novel structural variants of protein 4.1 are characterized. Both polypeptides are distinguished by aberrant electrophoretic mobilities on SDS-polyacrylamide gels, i.e., different from the normal apparent molecular weight of 78–80 kD (protein 4.1^{80}). Based on biochemical analysis in the preceding paper, both also have structural alterations in the functionally important spectrin-actin binding domain. One is a low molecular weight variant (protein $4.1^{68/65}$) that in heterozygotes causes elliptocytosis and moderately severe hemolysis. Chemical and enzymatic cleavage patterns, as well as immunologic data, revealed a deletion of part or all of the spectrin-acting binding domain in protein $4.1^{68/65}$. The second variant has a high molecular weight (protein 4.1^{95}), which in heterozygotes produces only mild elliptocytosis without anemia. Structural analysis of protein 4.1^{95} revealed an insertion of 15 kD adjacent to the spectrin-actin binding domain, and partial amino acid sequence data suggested a duplication of 57 amino acids (aa) of the 24-kD chymotryptic domain comprised part of the 15-kD insert.

In this paper we report cloning and sequencing of reticulocyte mRNA sequences encoding portions of proteins $4.1^{68/65}$ and 4.1^{95}, using reverse transcriptase and polymerase chain reaction (PCR) techniques (8) to amplify the likely rearranged regions as identified by the protein studies. This molecular characterization allowed us to deduce the primary structure of the abnormal 4.1 polypeptides. In addition, by comparison of the deletion and duplication end points in these mutant proteins with known exon-exon boundaries in that part of the protein 4.1 gene, we could deduce as well the genetic origin of these rearrangements. Finally, by comparing 4.1^{95} mRNA from reticulocytes vs lymphocytes, we were able to examine the effects of a gene rearrangement on splicing of an alternatively expressed exon of 21 amino acids in the spectrin-actin binding region.

Methods

Reagents. [^{32}P]dCTP and [^{32}P]γATP were obtained from Amersham Corp. (Arlington Heights, IL). Reverse transcriptase was supplied by Molecular Genetics Resources (Tampa, FL), RNasin by Promega Biotech (Madison, WI), Taq. polymerase by Perkin Elmer Cetus Corp. (Norwalk, CT), and restriction endonucleases by New England Biolabs

Address reprint requests to Dr. Mohandas, Cell and Molecular Biology Division, Building 74-157, Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, CA 94720.

Received for publication 23 January 1990.

1. *Abbreviations used in this paper:* aa, amino acid(s); HE, hereditary elliptocytosis; nt, nucleotide; PCR, polymerase chain reaction.

J. Clin. Invest.

0021-9738/90/08/0524/07 $2.00
Volume 86, August 1990, 524–530

(Beverly, MA). The plasmid vector Bluescript was purchased from Stratagene (San Diego, CA), and M13mp18 DNA from Pharmacia Fine Chemicals (Piscataway, NJ).

RNA isolation. Human reticulocyte mRNA was prepared as described previously (9) from blood of patients with elliptocytosis due to mutant protein 4.1. For normal protein 4.1 controls, reticulocyte mRNA was prepared from blood of sickle cell anemia patients undergoing exchange transfusion.

PCR amplification of protein 4.1 mRNA from reticulocytes. 1 μg of human reticulocyte poly (A)+ RNA was transcribed into single-stranded DNA at 42°C for 60 min in a 20-μl reaction containing: 40 mM KCl, 50 mM Tris-HCl (pH 8.3), 8 mM $MgCl_2$, 0.5 mM dNTPs, 50 pmol of antisense 4.1-specific oligonucleotide primer (see below), 20 U of AMV reverse transcriptase, and 10 U of RNasin. The reaction was diluted to 100 μl in Taq. polymerase buffer (50 mM KCl, 10 mM Tris-HCl, 1.5 mM $MgCl_2$, 0.1% gelatin) supplemented with 50 pmol of sense-strand oligonucleotide, additional dNTPs to a final concentration of 0.2 mM, and 5 U of Taq. polymerase. 30 cycles of amplification were performed using an automated Perkin Elmer Cetus thermal cycler under the following conditions: denaturation 30″ at 94°C; reannealing, 30″ at 50–65°C; extension, 1′45″ at 72°C. DNA fragments were analyzed by polyacrylamide gel electrophoresis. For sequencing, fragments were phosphorylated using polynucleotide kinase, blunt-end ligated into the Eco RV site of Bluescript and sequenced directly in the double-stranded plasmid. The oligonucleotides used in this paper are shown in Table I.

Results

Deduced structure of protein 4.1^{95} in reticulocytes. Normal protein 4.1^{80} consists of four structural domains of apparent molecular weights of 30, 16, 10, and 24 kD (10). In the preceding paper (7), biochemical analysis of these domains in the elongated variant protein 4.1^{95} revealed an insertion in the 10-kD spectrin-acting binding domain between Lys^{406} and Lys^{407}, which are the first two residues of this domain, and a known exon boundary. Partial amino acid sequence data suggested that the insert represents an internal duplication of at least 57 amino acids from Arg^{473}-Gln^{529} in the 24 kD domain (7) (Fig. 1 *A*). In order to characterize further this unusual structural variant, 4.1^{95} mRNA sequences were cloned and analyzed using 4.1-specific oligonucleotides (Table I) and PCR methodologies. Fig. 1 *B* presents the PCR strategy employed, depicting the map location of the expected amplification products relative to the structural domains of the protein, while Fig. 1 *C* shows a polyacrylamide gel analysis of the observed DNA bands produced in each reaction. Although the affected HE individuals are heterozygous for 4.1^{80} mRNA and 4.1^{95} mRNA, it was possible by judicious choice of oligonucleotides to preferentially amplify 4.1^{95} mRNA. Reaction I was designed to confirm the protein data suggesting that 24-kD sequences were inserted upstream of Lys^{407}, by using oligonucleotide primers flanking the putative 24-kD/10-kD junction to amplify the COOH-terminal end of the insert. No product was expected or observed when normal reticulocyte mRNA was used as a template, since the 5′ and 3′ primers are oriented away from each other (reaction Ia). However, due to the abnormal transposition of 24-kD sequences upstream of 10-kD sequences in 4.1^{95} mRNA, these oligonucleotides are oriented toward each other at the junction and can amplify a band of 130 bp. A band of appropriate size was indeed obtained from 4.1^{95} mRNA (reaction Ib). Sequence analysis of this DNA fragment confirmed that 4.1^{95} mRNA contains an insertion of 24-kD coding sequences upstream of Lys^{407} in the spectrin-actin binding domain.

Protein 4.1^{95} mRNA sequences encompassing the entire 15-kD insert were cloned next using a 5′ oligonucleotide located upstream in the 16 kD domain and a 3′ oligonucleotide precisely positioned at the 24-kD/10-kD junction at the 3′ end of the insert. Since the 3′ primer is derived from sequences in 24 kD and 10 kD that are discontinuous in 4.1^{80} mRNA, no product was obtained when normal reticulocyte mRNA was amplified (reaction IIa). However, a DNA band of 587 bp was obtained from 4.1^{95} mRNA (reaction IIb). Sequence analysis of this DNA revealed normal coding sequences in the 16-kD region up to residue Lys^{407} at the start of 10 kD, followed by an insert of 123 amino acids representing a duplication of an almost intact 10-kD region (Lys^{407}-Trp^{472}) plus residues Arg^{473}-Gln^{529} of 24 kD predicted from the protein studies.

The region of 4.1^{95} mRNA 3′ to the insert was cloned using oligonucleotide primers depicted in reaction III. Since the 5′ oligonucleotide is present within the duplicated portion of 24 kD and the 3′ primer is located in a downstream nonduplicated portionof 24 kD, two amplification products are possible using the 4.1^{95} mRNA template (reaction IIIb). Upon sequencing, the larger 502 bp band revealed the COOH-terminal end of the insert (i.e., the 24-kD/10-kD junction), followed by a normal 10-kD and 24-kD sequence. Normal 4.1^{80} mRNA possesses only one copy of the 5′ primer and thus yields only a single 133-bp band (reaction IIIa).

Table I. Sequences of Oligonucleotides Used for PCR Amplification of Protein 4.1 mRNA

Reaction	5′ (sense strand)	3′ (antisense strand)
I	5′-AATCCCAACCAAAGACGT-3′	5′-CTCCAACATTAAATTGCTAT-3′
II	5′-GCTGTCGATTCGGCAGACCGAAGTCCTCGGCCC-3′	5′-CTTTCTCTCTTTTTCTGGGC-3′
III	5′-AATCCCAACCAAAGACGT-3′	5′-CTCAGATGTGATAGTTTG-3′
IV	5′-GAAGAAAGCCCTCAATC-3′	5′-TGTAAAATTCCAAGGGACACCACGAACCTG-3′
V	5′-CCATTTGGGATAACGC-3′	5′-ACTTTTTGGCATTCTC-3′
VI	5′-GAAGAGAAGGTCATGGAACTGCATAAGTCA-3′	5′-GGGCCGAGGACTTCGGTCTGCCGAATCGACAGC-3′
VII	5′-CATCAAGGAGGCAAAGG-3′	5′-AAAGGATCAATGAGATCTAGAAGAAT-3′
VIII	5′-GCTGTCGATTCGGCAGACCGAAGTCCTCGGCCC-3′	5′-AGTGTGGACAATAGGGACGTCTTTGGTTGC-3′
IX	5′-GCTGTCGATTCGGCAGACCGAAGTCCTCGGCCC-3′	5′-CTTTCTCTCTTTTTCTGGGC-3′
X	5′-AATCCCAACCAAAGACGT-3′	5′-CTCAGATGTGATAGTTTG-3′

Note that the same 5′ primer was used in reactions II, VI and IX; also in reactions I, III, and X. The same 3′ primer was used in reactions II and IX. These sequence data are available from GenBank under accession number M14993.

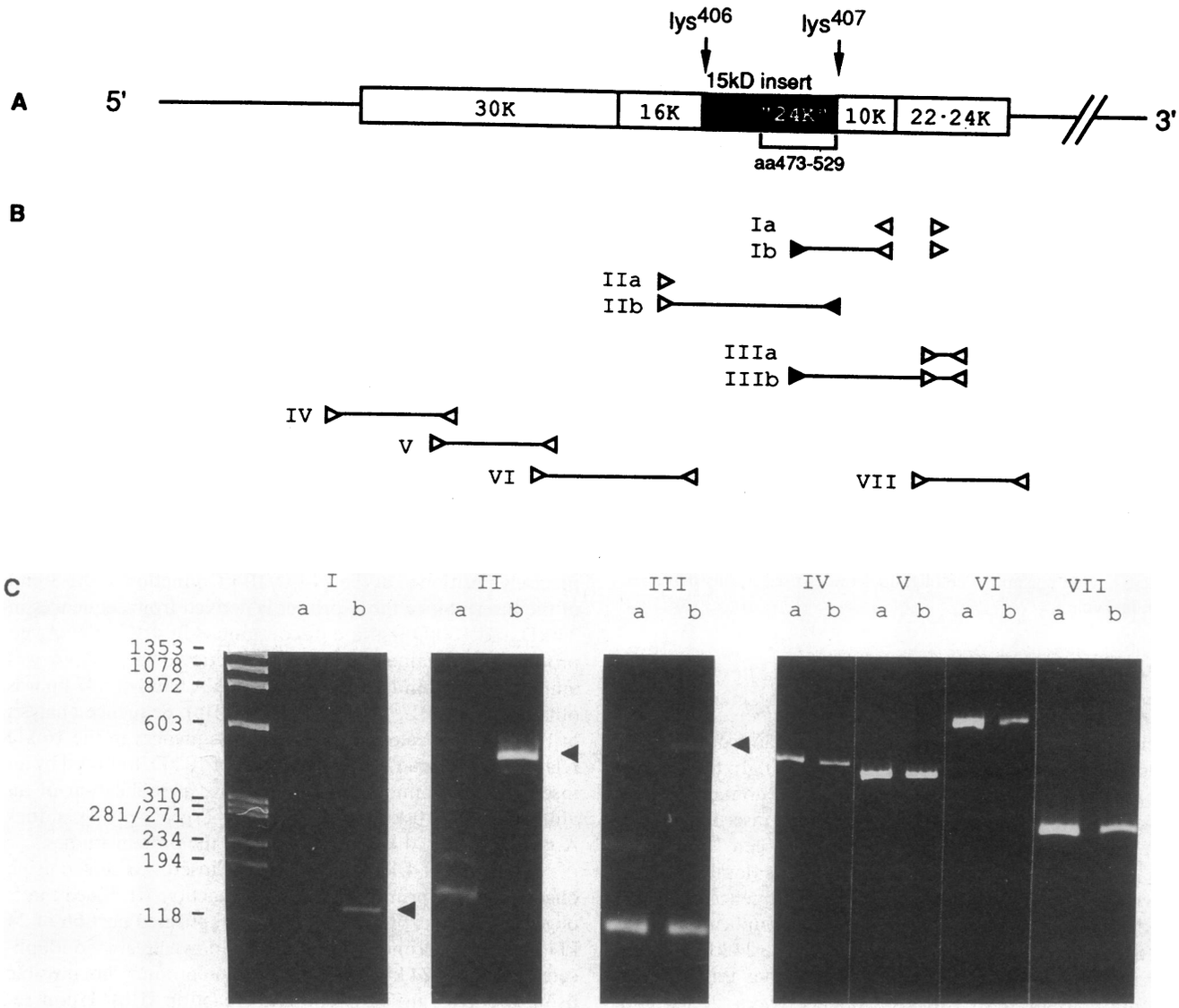

Figure 1. Scheme for 4.1^{95} mRNA analysis by PCR. (*A*) Structural model of 4.1^{95} mRNA. Narrow lines indicate 5′ and 3′ untranslated regions; open boxes indicate regions encoding the four structural domains of normal protein 4.1^{80}; black box shows the 15-kD insert between Lys^{406} and Lys^{407}, as deduced from protein studies. Apparent duplication of Arg^{473}-Gln^{529} from the 24-kD domain is indicated. (*B*) Amplified regions of 4.1^{95} mRNA. Black arrowheads indicate the location of oligonucleotides unique to the insert in 4.1^{95} mRNA but not found in normal 4.1^{80} mRNA; open arrowheads represent primers present in both 4.1^{80} and 4.1^{95} mRNAs. Reactions I-III analyze the 15-kD insert and its flanking sequence. Reactions IV-VII complete the characterization of upstream and downstream portions of 4.1^{95} mRNA. The notation "a" refers to 4.1^{80} mRNA and "b" indicates 4.1^{95} mRNA. (*C*) 5% polyacrylamide gel electrophoresis of amplified DNA products. Black arrows indicate the bands unique to 4.1^{95} mRNA.

Finally, amplification of sequences encoding the 30- and 16-kD domains, as well as the COOH-terminal end of 24 kD and parts of the 5′ and 3′ untranslated sequence, was accomplished using additional oligonucleotides (reactions IV–VII). All of these regions yielded PCR products that were indistinguishable in size from those obtained with normal 4.1^{80} mRNA substrates (compare lanes *a* and *b* in Fig. 1 *C*), suggesting that there are no other abnormalities or rearrangements present in 4.1^{95} mRNA.

Structure of protein $4.1^{65/68}$. A tentative structural model of the shortened protein $4.1^{65/68}$ polypeptide, based on protein analyses in the preceding paper (7), is shown in Fig. 2 *A*. To deduce the precise limits of the apparent deletion around the spectrin-acting binding region, the coding region of $4.1^{65/68}$ mRNA was amplified in a series of seven PCR reactions using 4.1-specific oligonucleotides. In each reaction, reticulocyte RNA from an individual with normal protein 4.1 was used as a control for comparison with the heterozygous HE patient. Reactions IV–VII yielded identical amplified DNAs in the normal and HE samples, indicating the absence of gross structural abnormalities in the 30-kD, 16-kD, and extreme COOH-terminal coding regions of $4.1^{65/68}$ mRNA. In contrast, the appearance of a unique shortened DNA band in reaction VIIIb (Fig. 2 *C, arrowhead*), in addition to the normal 539bp band

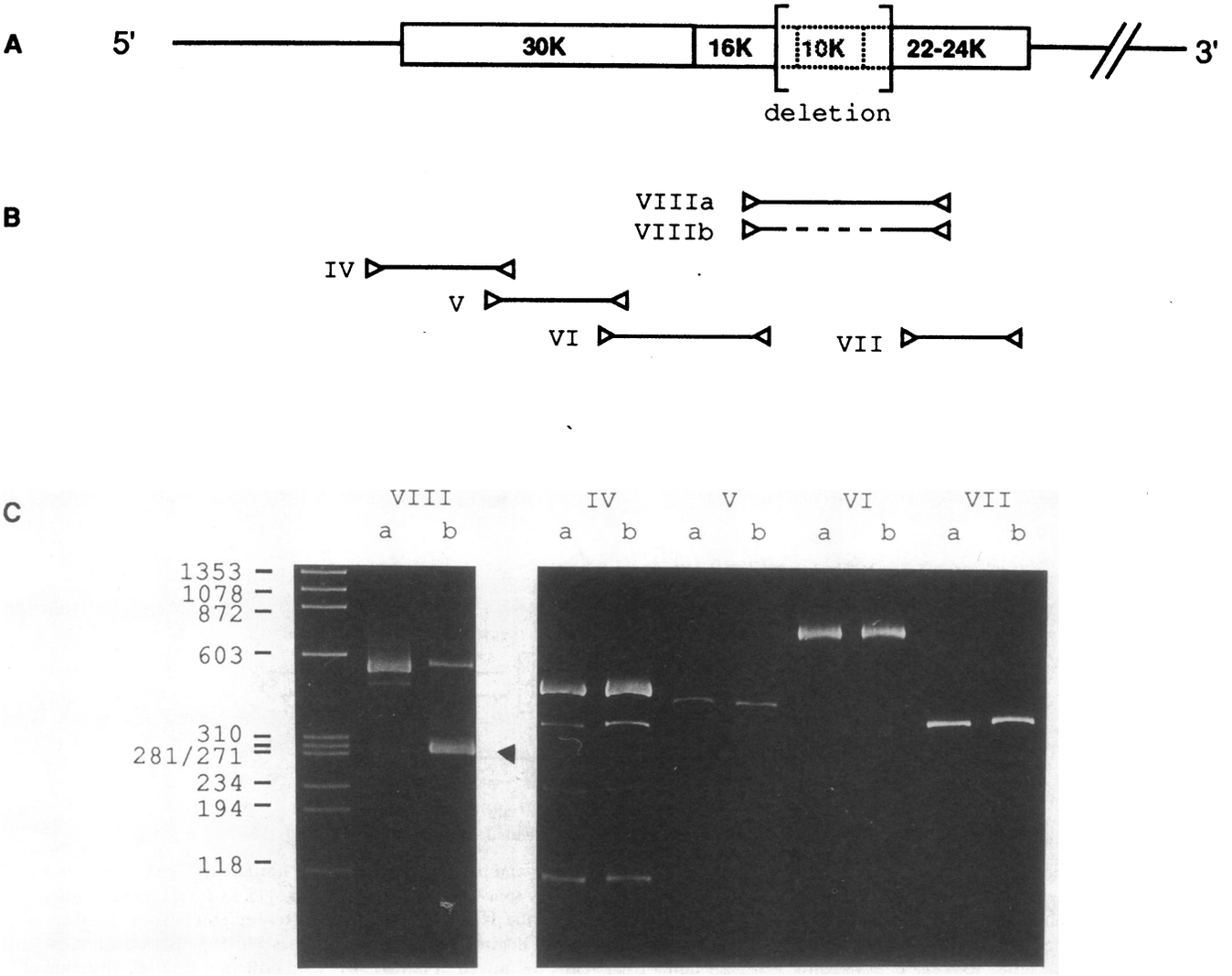

Figure 2. Scheme for $4.1^{65/68}$ mRNA analysis by PCR. (*A*) Structural model of $4.1^{65/68}$ mRNA. Narrow lines indicate 5′ and 3′ untranslated regions; open boxes indicate regions encoding the four structural domains of normal protein 4.1^{80}; brackets encompass region of possible deletion deduced from the protein studies. (*B*) Amplified regions of $4.1^{65/68}$ mRNA. Arrowheads indicate the oligonucleotide primer locations. Reactions VIII reveal a deletion in $4.1^{65/68}$ mRNA (dotted line in VIIIb) not seen in 4.1^{80} mRNA (solid line, VIIIa). The notation "a" refers to 4.1^{80} mRNA and "b" indicates $4.1^{65/68}$ mRNA. (*C*) 5% polyacrylamide gel electrophoresis of amplified DNA products. Black arrowhead indicates the band unique to $4.1^{65/68}$ mRNA.

(reaction VIIIa), reveals an abnormality in the 10-kD coding sequence. Nucleotide sequence analysis of the shortened DNA band defined a 240 nt deletion in $4.1^{65/68}$ mRNA that would encode a protein with a deletion of 80 aa from Lys^{407}-Gly^{486}. Included within this deletion is the entire 67-amino acid spectrin-actin binding domain (except for the first residue Lys^{406}) and a portion of the COOH-terminal 24-kD domain.

Genetic basis for protein 4.1 structural rearrangements. The deduced structures of proteins $4.1^{65/68}$ and 4.1^{95} contain internal deletions and duplications of peptides within the same small region of the protein 4.1 polypeptide, the functionally important spectrin-actin binding domain. To determine how these rearrangements in protein structure relate to 4.1 gene alterations, we compared the end points of these peptides to the deduced exon boundaries in this part of the gene. As described below and illustrated in Fig. 3, this comparison indicates that $4.1^{68/65}$ mRNA suffers a two-exon deletion, while 4.1^{95} mRNA possesses a three-exon duplication.

Exon boundaries in the spectrin-actin binding and 24-kD coding sequences have been deduced mainly through studies of alternative RNA splicing in normal protein 4.1^{80} mRNA (11). Multiple structural isoforms of erythroid 4.1 mRNA, characterized by insertion or deletion of nucleotide sequence cassettes corresponding to individual exons, arise by an alternative splicing mechanism (11, 12). Two of these alternatively expressed exons encode peptides Lys^{407}-Glu^{427} (21 aa) and Pro^{487}-Gln^{529} (43 aa), located exactly at the end points of the rearrangements in 4.1^{95} mRNA and $4.1^{68/65}$ mRNA (Fig. 3). To test whether the intervening 59 aa peptide Asp^{428}-Gly^{486} is encoded by a single exon, genomic DNA was amplified using oligonucleotides spanning this coding region. A single band of 177 nt was produced, indicating that this sequence is not in-

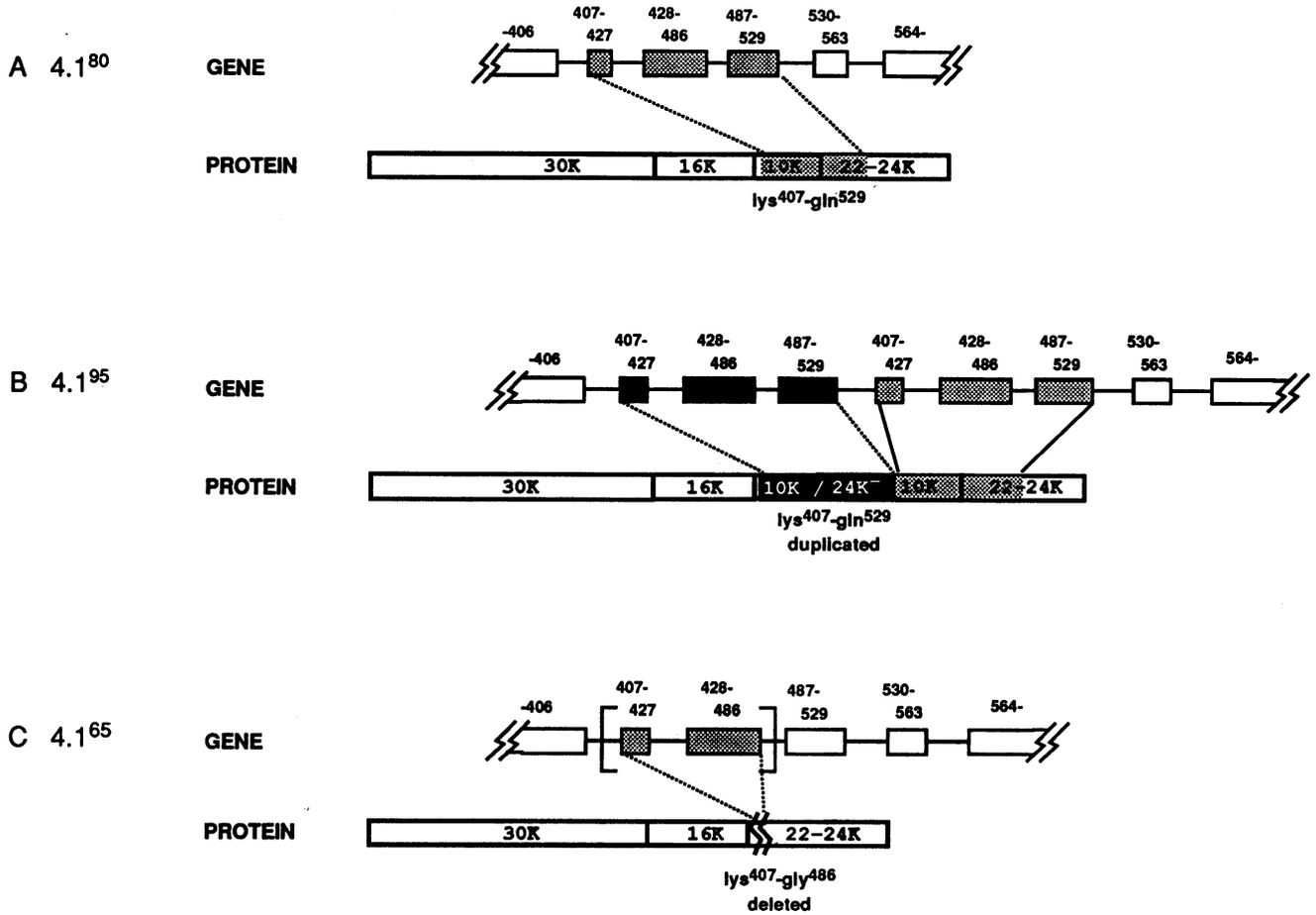

Figure 3. Genetic models of 4.1^{80}, 4.1^{95}, and $4.1^{65/68}$. (*A*) Structure of protein 4.1^{80} and partial gene map of the normal 4.1^{80} gene. The size and boundaries of four consecutive exons has been deduced by analysis of alternatively spliced 4.1 mRNA isoforms (11), by PCR amplification of genomic DNA (unpublished), and by sequencing of genomic DNA (in the case of the 34 amino acid exon). Boxes represent exons; numbers above each box indicate the amino acids encoded by that exon (reference 13, with the insertion of 34 aa, residues 530-563, that were not encoded in the original clone). Expression of the three exons encoding Lys^{407}-Gln^{529} is altered in protein 4.1^{95} and protein $4.1^{65/68}$. B. Structure of protein 4.1^{95} and proposed map of the 4.1^{95} gene. The duplicated sequence of Lys^{407}-Gln^{529} is proposed to be encoded by a duplication of three consecutive exons in the 4.1 gene. (*C*) Structure of protein $4.1^{65/68}$ and proposed map of the $4.1^{65/68}$ gene. The two shaded exons within the brackets are not represented in the shortened protein product.

terrupted by any introns and is thus a single exon (data not shown).

Alignment of the exon map with the domain map for protein-4.1 reveals that the 123-aa insertion in protein 4.1^{95} corresponds exactly to a duplication of three consecutive exons encoding 21, 59, and 43 aa (Fig. 3 *B*). Similarly, the 80-aa deletion in protein $4.1^{68/65}$ is due to deletion from the mRNA of only the first two of these exons (21 and 59 aa; Fig. 3 *C*).

Deduced structure of protein 4.1^{95} in lymphocytes. Comparison of erythroid (13) and lymphocyte (12) protein 4.1 cDNAs has revealed a tissue-specific splicing difference involving alternative splicing of the 21-aa exon. Most erythroid 4.1 isoforms express this 21 aa in their spectrin-actin binding domain, while the predominant lymphoid isoforms lack this peptide. In the 4.1^{95} gene, two copies of the 21-aa exon are present. The 5′ copy is apparently flanked by normal 5′ and 3′ neighboring exons, so its splicing behavior should be normal. The 3′ copy of this 21 aa exon, on the other hand, is flanked by a different 5′ exon than is usual. Its splicing pattern might be disrupted by this rearrangement.

To explore the alternative splicing of protein 4.1^{95} mRNA in lymphocytes, RNA was isolated from PHA-stimulated peripheral blood lymphocytes from an individual with 4.1^{95}. Lymphocyte 4.1^{95} mRNA was amplified with the oligonucleotides indicated in Fig. 4. Reaction IX spanned the 5′ copy of the 21-aa exon while reaction X encompassed sequences around the 3′ copy. In each reaction the predominant lymphoid product (lane *b*) was slightly smaller than the corresponding reticulocyte band (lane *a*). Sequence analysis showed that reticulocyte 4.1^{95} mRNA expressed both copies of the 21 aa exon, while the predominant lymphoid 4.1^{95} mRNA had both copies of this exon spliced out.

Discussion

We have characterized two interesting structural variants of erythroid protein 4.1 by direct analysis of the mutant polypeptides (7) and by molecular cloning of mutant mRNA sequences using PCR techniques (this report). In the case of

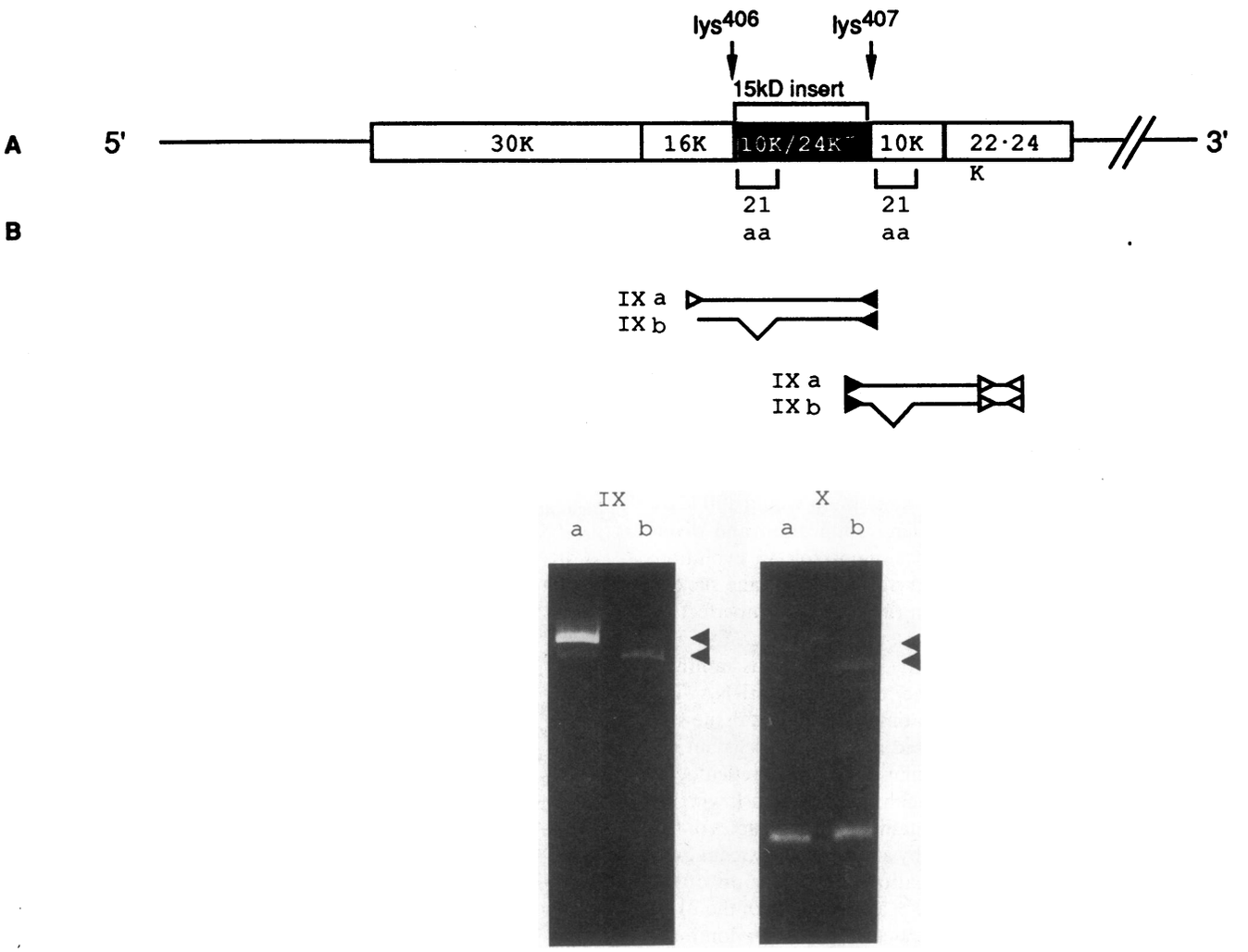

Figure 4. PCR analysis of splicing in reticulocyte and lymphocyte 4.1^{95} spectrin-actin binding domains. (*A*) Deduced structure of the coding region of 4.1^{95} mRNA in reticulocytes. Each of the two duplicated 10-kD domains contains a 21 aa peptide encoded by a 63 nt exon that is normally spliced out of lymphocyte 4.1^{80} mRNA (12). (*B*) Arrowheads show the location of oligonucleotides used to amplify separately the 10-kD domains in 4.1^{95} mRNA. Lines indicate deduced structures of the predominant mRNA isoforms produced from reticulocyte vs lymphocyte 4.1^{95} mRNA. The notation "a" refers to reticulocyte 4.1^{95} mRNA and "b" indicates lymphocyte 4.1^{95} mRNA. (*C*) Ethidium bromide stained 5% polyacrylamide gel electrophoresis of PCR products.

protein $4.1^{68/65}$, ektacytometric data showed that red cells from heterozygous patients expressing normal and shortened 4.1 exhibit mechanical properties similar to those from patients with 50% deficiency of 4.1, i.e., both have decreased fragmentation times. It is now clear that this phenotype results from an 80-aa deletion that removes the functionally important spectrin-actin binding domain. Protein 4.1^{95}, on the other hand, is elongated by virtue of a 123-aa duplication that results in expression of two spectrin-actin binding regions. At least one of these domains must remain in a functional conformation, since 4.1^{80+95} red cells are mechanically similar to normal cells.

The genetic mechanism(s) responsible for production of proteins with large internal structural rearrangements is of great interest. Probably the best documented mechanism for producing gene deletions and insertions is unequal crossing over between repetitive sequence elements in the DNA. For example, unequal crossing over between repeated sequences in the β-globin gene cluster produced reciprocal deletions and insertions manifested as unusual hybrid globin chains $\delta\beta$ (Lepore) and $\beta\delta$ (anti-Lepore) (14). While these globin rearrangements occurred over distances of $\sim$ 7 kb and involved separate genes, similar phenomena can occur within the confines of a single gene. Crossing over between 3.4-kb repeated sequences in the glycophorin C can delete either exon 2 (15) or exon 3 (16), producing shortened protein variants. Likewise, an internal duplication of 373 aa in a mutant LDL receptor has been attributed to unequal crossing over between middle repetitive Alu sequences in introns 1 and 8, duplicating 14 kb of the gene including exons 2-8 (17).

Our results show that sequences from two exons are missing from protein $4.1^{68/65}$, while three are duplicated within protein 4.1^{95}. By analogy with the examples cited above, these findings probably reflect gene rearrangements produced by unequal crossing over. The size of these putative duplications and deletions in the protein 4.1 gene are probably > 10 kb, based on preliminary Southern blotting and genomic cloning exper-

iments (not shown). Both putative 4.1 gene rearrangements occur in the same region of the gene; however, they cannot represent reciprocal products of a single crossing over event, since they possess distinct 3′ boundaries. Rather, it may be that this portion of the protein 4.1 gene is a hot spot for such recombination events. Previous analyses of numerous deletions in the Duchenne muscular dystrophy gene have demonstrated precedence for recombination hot spots (18–20).

One functional contrast exists between the LDL receptor mutant and 4.1^{95}. Whereas the elongated LDL receptor is functionally defective, protein 4.1^{95} apparently retains its ability to interact with other membrane skeletal elements, based on ektacytometric data showing that $4.1^{80}/4.1^{95}$ cells have nearly normal mechanical properties (7). It is interesting to speculate on the evolutionary possibilities of such a protein with a duplicated functional region. This unusual protein 4.1^{95} might be in an advantageous situation in which to evolve new binding interactions or new mechanisms to regulate existing interactions, since alterations in one domain would still leave one functional domain intact. Gene duplication and divergence events are known to play a major role in evolution; many genes, including even the erythroid membrane skeletal proteins spectrin (21) and ankyrin (22), contain imperfect repeats indicative of past duplication events.

The putative duplication in the 4.1^{95} gene has ramifications with respect to splicing of the resulting pre-mRNA. The 21-aa exon is known to be differentially spliced in a tissue-specific pattern such that it is expressed in erythroid cells but not in lymphocytes (13, 12). The availability of both reticulocyte and lymphocyte 4.1^{95} mRNA made it possible to ask whether this differential expression is maintained for both copies of this exon, since the 3′ copy is flanked by a different upstream exon than is usual. By PCR amplification of this region of the mRNA it was determined that the 5′ and 3′ copies of the 21-aa exon behave similarly: both are spliced *into* the predominant erythroid 4.1 mRNA isoforms and *out of* the most abundant lymphoid isoforms. This finding implies that the signals for alternative mRNA processing of the 21-aa exon must be a function of the sequence in or near this exon itself, independent of the neighboring exon(s).

Acknowledgments

We would like to acknowledge James Harris for his expert assistance in preparing this manuscript. This work was supported in part by grants from the National Institutes of Health (DK-32094) and in part by the Office of Health and Environmental Research, Division of the U.S. Department of Energy under contract No. DE-AC03-765F00098.

References

1. Tchernia, G., N. Mohandas, and S. B. Shohet. 1981. Deficiency of skeletal membrane protein 4.1 in homozygous hereditary elliptocytosis. *J. Clin. Invest.* 68:454–460.

2. Alloisio, N., L. Morle, E. Dorleac, O. Gentilhomme, D. Bachir, D. Guetarni, P. Colonna, B. M., Z. Zouaoui, L. Roda, D. Roussel, and J. Delauney. 1985. The heterozygous form of 4.1 (−) hereditary elliptocytosis [the 4.1 (−) trait]. *Blood.* 65:46–51.

3. Conboy, J. G., N. Mohandas, G. Tchernia, and Y. W. Kan. 1986. Molecular basis of hereditary elliptocytosis due to protein 4.1 deficiency. *N. Engl. J. Med.* 315:680–685.

4. McGuire, M., B. L. Smith, and P. Agre. 1988. Distinct variants of erythrocyte protein 4.1 inherited in linkage with elliptocytosis and Rh type in three white families. *Blood.* 72:287–293.

5. Lovric, V. A., R. J. Walsh, and M. A. Bradley. 1965. Hereditary elliptocytosis: genetic linkage with the Rh chromosome. *Aust. Ann. Med.* 14:162–166.

6. Lambert, S., J. Conboy, and S. Zail. 1988. A molecular study of heterozygous protein 4.1 deficiency in hereditary elliptocytosis. *Blood.* 72:1926–1929.

7. Marchesi, S., J. Conboy, P. Agre, J. T. Letsinger, V. T. Marchesi, D. W. Speicher, and N. Mohandas. 1990. Molecular analysis of insertion/deletion mutations in protein 4.1 in elliptocytosis. I. Biochemical identification of rearrangements in the spectrin/actin binding domain and functional characterizations. *J. Clin. Invest.* 86:516–523.

8. Scharf, S. J., G. T. Horn, and H. A. Erlich. 1986. Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science (Wash. DC).* 233:1076–1078.

9. Temple, G. F., J. C. Chang, and Y. W. Kan. 1977. Authentic β-globin mRNA sequences in homozygous β^0 thalassemia. *Proc. Natl. Acad. Sci. USA* 74:3047–3051.

10. Leto, T. L., and V. T. Marchesi. 1984. A structural model of human erythrocyte protein 4.1. *J. Biol. Chem.* 259:4603–4608.

11. Conboy, J. G., J. Chan, N. Mohandas, and Y. W. Kan. 1988. Multiple protein 4.1 isoforms produced by alternative splicing in human erythroid cells. *Proc. Natl. Acad. Sci. USA.* 85:9062–9065.

12. Tang, T. K., T. L. Leto, I. Correas, M. A. Alonso, V. T. Marchesi, and E. J. Benz. 1988. Selective expression of an erythroid-specific isoform of protein 4.1. *Proc. Natl. Acad. Sci. USA.* 85:3713–3717.

13. Conboy, J. G., Y. W. Kan, and N. Mohandas. 1986. Molecular cloning of protein 4.1, a major structural element of the human erythrocyte membrane skeleton. *Proc. Natl. Acad. Sci. USA.* 83:9512–9516.

14. Weatherall, D. J., and J. B. Clegg. 1981. The Thalassemia Syndromes. 3rd ed. Blackwell Scientific Publications, Oxford.

15. Chang, S., M. E. Reid, J. Conboy, Y. W. Kan, and N. Mohandas. 1989. Molecular characterization of erythrocyte glycophorin C variants. *Blood.* 74(Suppl. 1):26*a*. (Abstr.)

16. Colin, Y., C. Le Van Kim, A. Tsapis, M. Clerget, L. d'Auriol, J. London, F. Galibert, and J.-P. Cartron. 1989. Human erythrocyte glycophorin C. *J. Biol. Chem.* 264:3773–3780.

17. Lehrman, M. A., J. L. Goldstein, D. W. Russell, and M. S. Brown. 1987. Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. *Cell.* 48:827–835.

18. Wapenaar, M. C., T. Kievits, K. A. Hart, S. Abbs, L. A. Blonden, L. A. J. den Dunnen, P. M. Grootscholten, E. Bakker, C. Verellen-Dumoulin, M. Bobrow, G. J. B. van Ommen, and P. L. Pearson. 1988. A deletion hot spot in the Duchenne muscular dystrophy gene. *Genomics.* 2:101–108.

19. Forrest, S. M., G. S. Cross, T. Flint, A. Speer, K. J. H. Robson, and K. E. Davies. 1988. Further studies of gene deletions that cause Duchenne and Becker muscular dystrophies. *Genomics.* 2:109–114.

20. Grimm, T., B. Muller, M. Drier, E. Kind, T. Bettecken, G. Meng, and C. R. Muller. 1989. Hot spot of recombination within DXS164 in the Duchenne muscular dystrophy gene. *Am. J. Hum. Genet.* 45:368–372.

21. Speicher, D. W., and V. T. Marchesi. 1984. Erythrocyte spectrin is composed of many homologous triple helical segments. *Nature (Lond.).* 311:177–180.

22. Lambert, S., J. Lawler, H. I. Yu, D. W. Speicher, J. T. Prchal, and J. Palek. 1988. A cDNA for human erythrocyte ankyrin. *Blood.* 72(Suppl. 1):30*a*. (Abstr.)