TRANSACTIONS OF THE AMERICAN CLINICAL CLIMATOLOGICAL ASSCOCIATION, VOL. 108, 1997

ROLE OF TISSUE SPECIFIC ALTERNATIVE PRE-mRNA SPLICING IN THE DIFFERENTIATION OF THE ERYTHROCYTE MEMBRANE

EDWARD J. BENZ, JR. and SHU CHING HUANG

BALTIMORE, MD

INTRODUCTION AND BACKGROUND

The body contains many distinct cell types, each possessing specialized size, shape, and functions, even though each inherits the same complement of genes. This is possible because each cell expresses only a few of the thousands of genes present in its genome. These restricted patterns of gene expression are strictly controlled according to cell type, developmental stage, differentiation state, and changes in the environment. As indicated by Figures 1 and 2, genetic information is expressed by copying the DNA nucleotide base sequence (gene) into an initial RNA transcript called pre-messenger RNA, intranuclear processing of pre-mRNA into mature mRNA, and its translation in the cytoplasm into the amino acid sequence of the protein encoded by that gene (cf Reference 1 for review). The term "regulation of gene expression" is used to refer to processes that determine which genes are expressed in a given cell, and the timing, extent, and inducibility of expression. Gene regulation is essential for normal ontogeny, differentiation, cell growth and death, and homeostatic responses to changing conditions. The existence of complex tissues and organs is impossible without mechanisms that provide for restricted but plastic patterns of gene expression in individual cells.

Much evidence suggests that the first step in gene expression, copying of the DNA sequence into the pre-mRNA precursor, is critical for the regulation of most genes. Most human DNA is effectively sequestered from the transcriptional apparatus by condensed chromatin structure. Genes must be "activated" for expression by recognition of specific flanking sequences that mark them as appropriate targets for expression. These regulatory sequences (promoters, enhancers, silencers) alter the ability of the genes within or near which they reside to be transcribed by interacting with proteinaceous "transcription factors"

From the Departments of Medicine and Molecular Biology and Genetics, The Johns Hopkins University School of Medicine.

Address for reprints: Edward J. Benz, Jr., M.D., Department of Medicine, Johns Hopkins Hospital, Suite 9028, 1830 E. Monument Street, Baltimore, MD 21205.

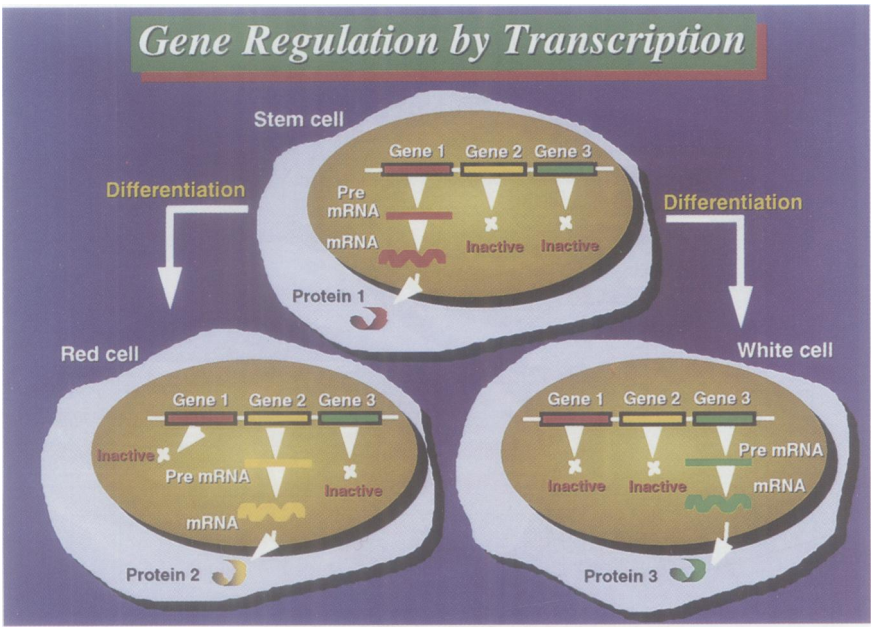


FIG. 1. Gene Regulation by Selective Transcription. The diagram shows an idealized stem cell in which three genes have the potential to be transcribed into mRNA, resulting in potential combinations of none, any one, any two, or all three of the encoded proteins. The diagram depicts the simple situation wherein transcription of Gene 1 maintains the stem cell state, while repression of Gene 1 and transcription of Gene 2 occurs during red cell differentiation and repression of Gene 1 but transcription of Gene 3 leads to white cell differentiation. This is the type of selective gene expression most observers think of when the term gene regulation is used.

that bind to their target sequences to promote or repress transcription. Knowledge in this area is advancing rapidly, but no unifying view that explains the actual basis for gene activation and repression has emerged.

Gene regulation by transcription alone is appealing in its simplicity. As shown in Figure 1, the distinct types of proteins produced in two different types of blood cells could be explained solely on the basis of which genes were or were not transcribed into mRNA. The pattern of gene expression could then be inferred simply by straightforward measurement of the mRNAs present in each cell. However, many genes are now known to be regulated at post transcriptional as well as transcriptional levels. Well known examples include the regulation of ferritin mRNA levels by iron-dependent alterations in stability, and the selective stabilization or destabilization of mRNAs encoding growth-regulating genes in responses to growth-altering conditions.

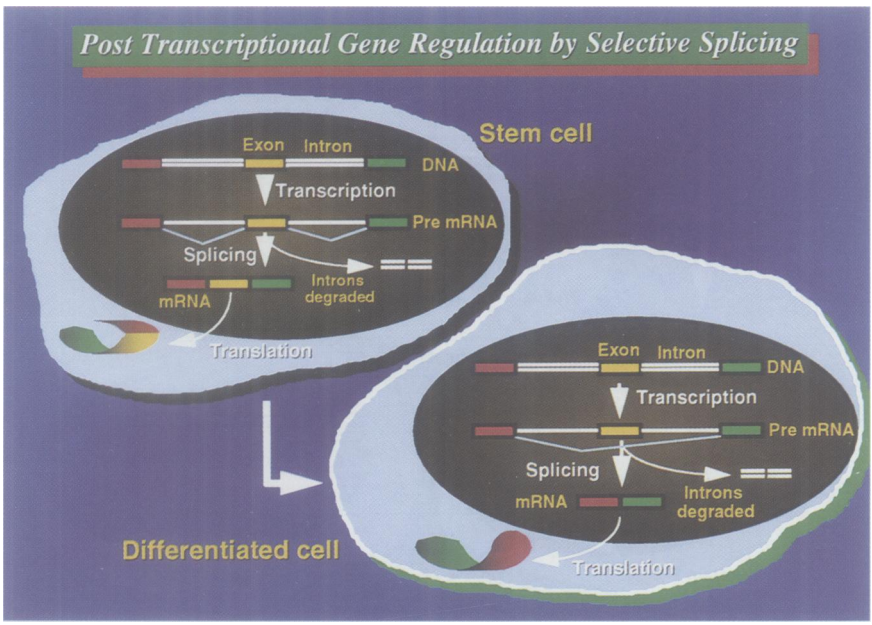


FIG. 2. Regulation of Gene Expression by Alternative pre-mRNA Splicing. The diagram shows a single gene containing three exons, red, yellow, and green. In the idealized stem cell, all of the splicing sites (see text) are used, resulting in slicing of all three exons into mature mRNA. (The splicing pattern is indicated by the thin pale blue lines.) The resulting protein has the three domains indicated by the colored symbol. In the differentiated cell, the splice sites surrounding the central (yellow) exon are not recognized, and, as indicated by the thin blue line, the exon is "skipped" or omitted from mature mRNA. The encoded protein is identical to the one expressed in the stem cell except that it is missing the domain in the center of the protein. The "differentiated" product is thus likely to have similar properties, but to be distinct in significant details from the stem cell isoform.

A particularly complicated, important form of gene regulation is the regulated splicing of pre-mRNA into mRNA along alternative pathways in different tissues, at different developmental and differentiation stages, or in response to changing conditions. *Alternative pre-mRNA splicing* permits a single gene to encode several distinct forms of a protein that differ in structure, binding properties, intracellular localization, or stability (Figure 2). Regulated splicing thus provides a means for generating multiple products from a single gene that can provide either qualitative (functionally altered forms) or quantitative (products exhibiting altered translatability and/or stability) adaptation to varying physiologic demands. This greatly increases the flexibility and complexity of the genetic regulation actually occurring in most cells. In this paper, we consider the basic features of this funda-

mental element of eukaryotic gene expression, and illustrate its impact on erythrocyte differentiation by an example, Protein 4.1 gene expression, that we have encountered in our own studies.

BASIC FEATURES OF PRE-mRNA SPLICING

Genes consist of a linear array of blocks of sequence that ultimately appear in mature mRNA (exons), interrupted by introns, which are blocks of sequence that are excised from the pre-mRNA during splicing in the nucleus. "Pre-mRNA" is the initial transcript of a gene, the precursor of mature mRNA (Figure 2). It is a copy of all of the exons and introns of the entire DNA gene. Pre-mRNA is extensively processed in the nucleus to yield an mRNA moiety that contains only exons, and is transported to the cytoplasm. The process by which introns are removed and the remaining exons precisely spliced together is called mRNA splicing (2–4).

Splicing occurs on intranuclear ribonucleoprotein structures called snRNP's or spliceosomes (cf. 2–4). They consist of small nuclear RNA's (U1, U2, etc.) bound to structural and RNA-binding proteins. These complexes recognize exon-intron boundaries in the pre-mRNA, hold these RNA regions in the proper configuration for splicing, and insure precise cleavage of introns and reunion of the exons. They are susceptible to modulation by reversibly bound "splicing factors".

Key sequences located at or near the exon-intron boundaries in pre-mRNA (*cis* elements) are recognized and attacked by snRNP's (2). At the 3′ end of the intron is an "acceptor site", or 3′ splicing site; it has the invariant dinucleotide "AG" which must be present for splicing to occur. The AG cleavage site is surrounded by the *consensus* sequence: **YAG/G** (Y = pyrimidine), and is preceded in the intron by a 14–20 base pyrimidine-rich sequence, usually located within the first 20–35 bases upstream of the exon (polypyrimidine tract). At the 5′ boundary of the intron is the "donor" site or 5′ splice site; the invariant dinucleotide "GT" is the cleavage and reunion site. It is surrounded by the consensus sequence **AG/*GT*RAGT** (R, purine). A third key site usually located 20–60 bases upstream of the exon is the branch-point "lariat" site which is critical for the formation of a looped RNA intermediate (2–4).

The vast majority of exon-intron boundaries in most genes match the consensus sequences closely. Pre-mRNA transcripts of these regions thus have a high affinity for snRNP's and "constitutive" splicing factors found in virtually all cells. These pre-mRNA's tend always to be spliced

precisely and efficiently in any tissue in which the gene is transcribed. This behavior is called " constitutive" splicing.

"Alternative mRNA splicing" is the term used to describe the phenomenon whereby different mature mRNAs are generated from a single pre-mRNA species in different tissues by regulated use or nonuse of splicing signals surrounding a given exon. About 5% of genes are alternatively spliced (4–6). They usually contain one or more exons bounded by "weak" splice sites (poor consensus match) or other sequence anomalies nearby that cause the exons to interact less well with generic factors and spliceosomes. "Weaker" sites deviate from consensus, have a lesser affinity for snRNPs, and, presumably, require additional "tissue-specific" factors or factor modifications in order to be recognized. They can thus be differentially included in or excluded from mature mRNA, depending on the array of factors active in a particular cell type, developmental stage, physiologic state, or stage of differentiation. In some cases, sequences in the intron (e.g. long pyrimidine tracts or short purine "blocks") or even within the exon itself can also act as "enhancers" or "repressors" of splicing. These sequences presumably interact with RNA or protein factors that favor factor recognition and splicing of the exon, or block the use of that splice site.

Alternative RNA splicing need not always involve complete inclusion or skipping of an entire exon. In some cases, two or more alternative boundaries at one end (or both) of an exon can exist. The "weaker" of the alternative splice boundaries in these cases is often called a "cryptic" donor or acceptor site because it tends to be utilized only in the minority of cell types or biological states. Mutations that create cryptic splice sites or cause them to resemble strong sites more closely are well documented causes of many inherited disorders, e.g., thalassemias and hereditary elliptocytosis.. These mutations lead to missplicing of the pre-mRNA, resulting in a non translatable mRNA or one that encodes an anomalous protein product (1).

It is curious that mutations abolishing a normal splice site invariably cause utilization of otherwise silent cryptic sites, rather than leading to accumulation of unspliced pre-mRNA. This suggests that alternative splicing may depend in part on *competition* between weaker and stronger sites, rather than absolute ability or inability to interact with constitutive splicing factors (2–6). Indeed, some alternate splicing pathways appear to be susceptible to modulation by *quantitative* variations in constitutive factor levels rather than absolute presence or absence of a "tissue-specific" factor. This reality has greatly impeded the search for tissue-specific factors that account for well characterized examples of highly regulated splicing. The mecha-

nisms controlling alternative splicing are now under active study in many laboratories, including our own, but will not be considered further in this review.

BIOLOGICAL IMPLICATIONS OF ALTERNATIVE SPLICING

Alternative splicing clearly provides cells with additional levels of genetic control and flexibility (2,5). A single pre-mRNA can be spliced differentially according to gender, cell type, physiologic state, or developmental stage to produce multiple mRNAs and related but distinct proteins. Thus, as illustrated by Figure 2, a gene encoding 3 exons, one of which is spliced into mRNA in the stem cell, but omitted from mature mRNA in the differentiated progeny of that cell, can encode two proteins, differing only by the presence (in this example, in the stem cell) of the amino acids encoded by the exon selectively included in the stem cell. These two protein isoforms could then serve modified purposes in the stem cell and its differentiated progeny by virtue, for example, of different trafficking behaviors. One example of this added diversity and flexibility is the appearance of distinct membrane-bound and secreted forms of IgM during immunocyte differentiation; these forms differ because regulated pre-mRNA splicing events include or exclude exons encoding protein export signals. A single gene can thus code for either or both the membrane-bound and exported forms of the same immunoglobulin species, depending on differentiation stage. Indeed, large numbers of isoforms can be produced from a single gene if multiple exons are alternatively spliced; thus, two alternately spliced exons provide for four potential products: three, eight; four, sixteen, etc. If n is the number of selectively spliced exons, 2^n is the potential number of products that could arise by combinatorial retention or omission of each exon.

Alternative splicing provides powerful ways to allow a gene to code for many modular "variations on a theme" of its protein product. A protein possessing a basic structural property or catalytic activity could then be "customized" for a particular cell type or state by the splicing in or splicing out of peptide domains that modify its localization, affinity for subunits or co-factors, susceptibility to phosphorylation, etc. Many examples have been encountered (2–6), including bound and soluble forms of receptors, growth hormone and calcitonin isoforms with distinct properties, and the *dsx* and *tra gene* products in *Drosophila*, whose alternative splicing patterns mediate the determination of gender in early embryos (7).

In addition to allowing a diversity of protein properties to be encoded by a single gene, alternative splicing can also permit the gene to provide for variable *amounts* of protein product in different cell types or physiologic states. Differential splicing pathways that alter the 5′ or 3′ untranslated regions of the mature mRNA in different cell types are well documented. Since these regions are key determinants of cytoplasmic mRNA stability and translational efficiency, the net effect is to produce mRNA's accumulating in different amounts in different tissues without altering the transcription of the gene, or mRNA's yielding different amounts of translated protein product without altering the mRNA level. Finally, domains affecting the proteolytic sensitivity of the translated product can be included or excluded in the protein by alternative splicing, thus altering the turnover time and steady state amounts of the protein (2–6).

The apparent advantages of alternative splicing seem clear. A single gene can maximize its coding range and control accumulation of its product without the need to duplicate or "re-invent" useful coding sequences that have evolved. A great deal of control can be packaged into smaller amounts of DNA. Indeed, the incredibly compact genomes of viruses make extensive use of alternate splicing (examples include the key *rev* and *tat* products of *HIV-1*, and the large and small T antigens of *SV40* virus) (1–6). Unfortunately, the known behavior of many genes pre-empts any teleological explanation for the existence of alternative splicing or for the fact that transcriptional regulation controls some genes, while others employ splicing control, and still others both. Equally mysterious is the use of gene duplication and mutational drift to generate isoform diversity for some protein families, while others arise by means of tissue-specific splicing. For example, the human globins and the isoforms of Na, K-ATPase, which exhibit only subtle variations in behavior, are all encoded by independent genetic loci arising from a primordial gene by duplication (1,8).

Regardless of its theoretical basis, regulated alternative splicing is clearly germaine to molecular medicine. Many important normal and pathologic phenomena are governed by splicing regulation. This point is, perhaps, best reinforced by considering the impact of alternative mRNA splicing of Protein 4.1 pre-mRNA on the biogenesis of the red cell membrane. Protein 4.1 is a particularly striking and instructive example of the enormous diversity of proteins and, potentially, protein functions that can be generated from a single gene by this regulated splicing mechanism.

PROTEIN 4.1 - A CYTOSKELETAL PROTEIN FAMILY GOVERNED BY ALTERNATIVE SPLICING

The erythrocyte cytoskeleton (Figure 3) consists of spectrin oligomer and short actin fibers organized into a hexagonal lattice bound to protein 4.1 and other proteins at "junctional complexes". This firm, flexible structure is attached to the lipid bilayer by at least two linkages: [1] binding of protein 4.1 to the cytoplasmic domains of embedded transmembrane proteins such as glycophorin C and band 3; and, [2] attachment of band 3 to spectrin by ankyrin (9–11). This basic organization provides for the strength, flexibility, and resiliency that characterize the viscoelastic properties of the red cell membrane. Inherited defects of the major cytoskeletal proteins are associated with hereditary spherocytosis (HS), hereditary elliptocytosis (HE), and hereditary pyropoikilocytosis (HPP). Each condition is characterized by increased red cell fragility, abnormal red cell shape, and reduced resistance to shear stress (9,12). The protein defects associated with these anemias have been shown to derange the cytoskeletal interactions shown in Figure 3 (9–12). Erythrocyte protein 4.1 is an abundant (200,000 copies/red cell) (9,10) and critical element of this basic structure. It

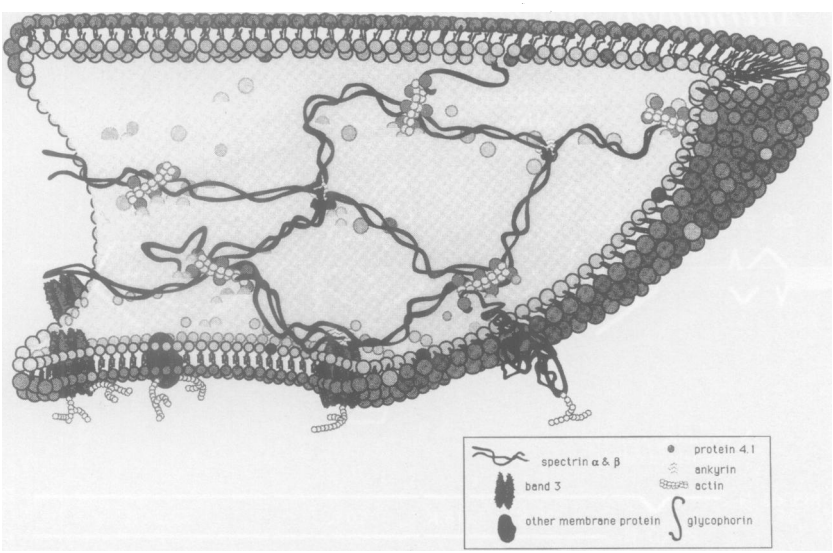


FIG. 3. Gene Regulation by Selective Transcription. See text for details. The meshwork formed by the α and β subunits of spectrin is shown linked into a junctional complex by actin and protein 4.1. Points of membrane attachment mediated by ankyrin-band 3 and protein 4.1-glycophorin C are shown. Not shown are lower affinity protein 4.1-band 3 linkages, and protein 4.1-glycophorin C interactions that also participate in the complexes. (Adapted from Reference 1).

stabilizes the "horizontal" interactions of spectrin and actin and also provides a key site for indirect "vertical" attachment to the lipid bilayer. Its basic structure is illustrated by Figure 4. The 80 KD erythroid phosphoprotein can be subdivided into four fragments (domains) by limited chymotryptic digestion (10,13,14). The amino-terminal (30 KD) domain is thought to contain the binding sites for integral membrane proteins. The 10 KD SAB domain contains the amino acids needed to bind to spectrin/actin (13,14). No specific functions have yet been assigned to the 16 KD region (10, 12). Our preliminary results (data not shown) suggest that the 22-24 KD domain interacts with DNA methyl transferase and nuclear and cytoplasmic matrix proteins such as NuMA and elongation factor 1α.

Erythroid and Nonerythroid 4.1 Forms Arise from Alternative mRNA Splicing

Isoforms of protein 4.1 are present in many cell types, within which they are heterogeneous with regard to molecular weight, abundance,

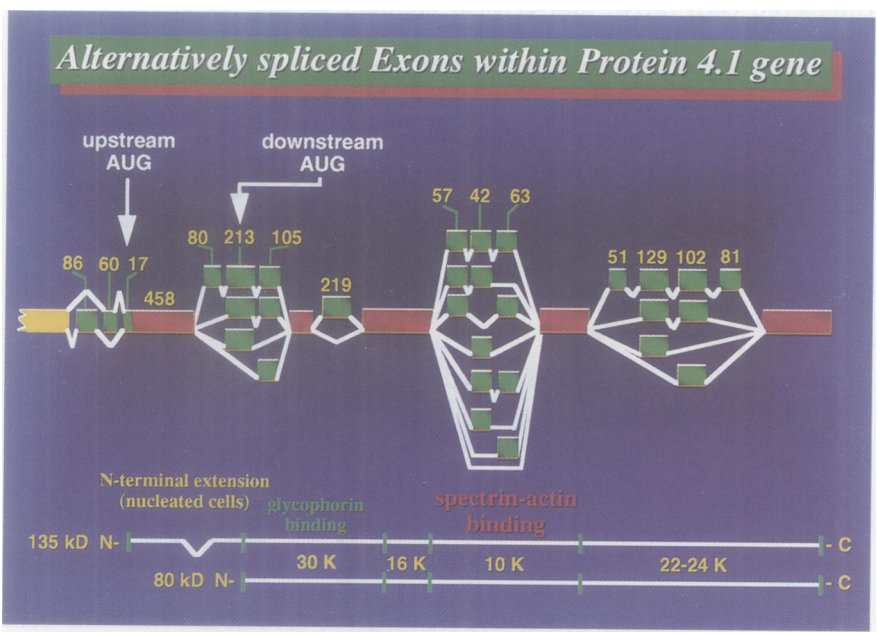


FIG. 4. Alternatively Spliced Exons Within Protein 4.1 Gene. Detailed exon/intron organization is not shown (cf. Figure 7). Only the alternatively-spliced exons (green boxes) and constitutive exons (red boxes) are shown. The figure indicates the regions within which splicing alters the amino terminus and the SAB. (Adapted from Reference 1).

and cellular localization (15–19). Nonerythroid forms are concentrated at points of cell-cell or cell/matrix contact, along stress fibers, on cytoplasmic matrix-like structures, in centrioles, in the perinuclear region, and, possibly, in the nucleus. The basic importance of these proteins to cell structure and function is implied by their prevalence and homology in many species.

We and others have demonstrated that these multiple isoforms arise from a *single* gene by complex patterns of tissue and differentiation stage-specific alternative mRNA splicing (20–23) of the initial pre-mRNA transcript. The human gene resides on chromosome 1p33–p34.2 near the Rh locus (24); it is very large, encompassing 250–350 kb (25). At least fourteen of the 24 known exons exhibit alternative mRNA splicing (Figure 4). A bewilderingly large array (potentially 2^{14}) of isoforms could conceivably arise from combinatorial splicing of these exons. Fortunately, we have discovered two splicing events involving selection of a limited number of exons that seem to have profound effects on the differentiation of erythroid and non-erythroid cytoskeletal structures. These involve the exons encoding the 10KD domain that governs spectrin actin binding, and the 5′ extremity of the mRNA. The latter results in isoforms that vary dramatically in size and intracellular localization.

1) Alternative Splicing Within the 10 KD Domain (Spectrin Actin Binding Domain)

Protein 4.1 pre-mRNA is striking for the complexity of its alternate splicing pathways. These have been most thoroughly characterized for exons 14, 15, and 16. These three exons within the SAB contain, in the 5′ to 3′ direction, 57, 42, and 63 nucleotides, and encode three peptide motifs: N-19-14-21-C, where 19, 14, and 21 refer to the number of amino acids encoded by each exon (Figure 5). Eight combinatorial splicing products could be generated within this small region alone. However, we have encountered only three of these products in meaningful amounts in cells. We and others (19–23) have shown that regulation of the splicing pathways for exons 14-15-16 is especially striking during erythroid differentiation. Except for brain, in which mRNA containing all three of these exons (+++) is reasonably abundant, the predominant mature mRNA species produced in non-erythroid cells and uninduced mouse erythroleukemia cells (MELC) contains none of the three exons (—). In maturing erythroblasts and after induction of erythroid maturation in MELC, however, a striking switch in the splicing pathway occurs so that mRNA containing only the 63 bp

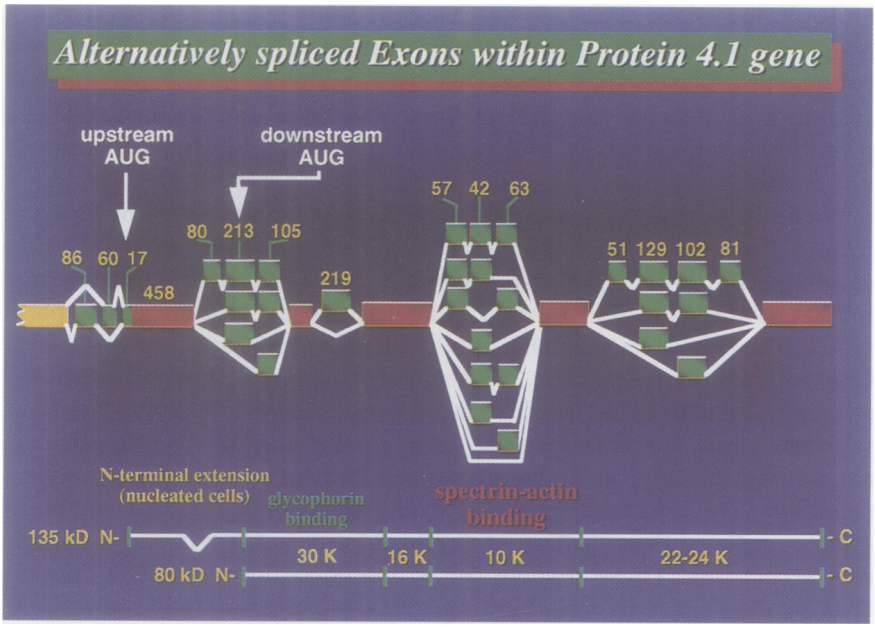
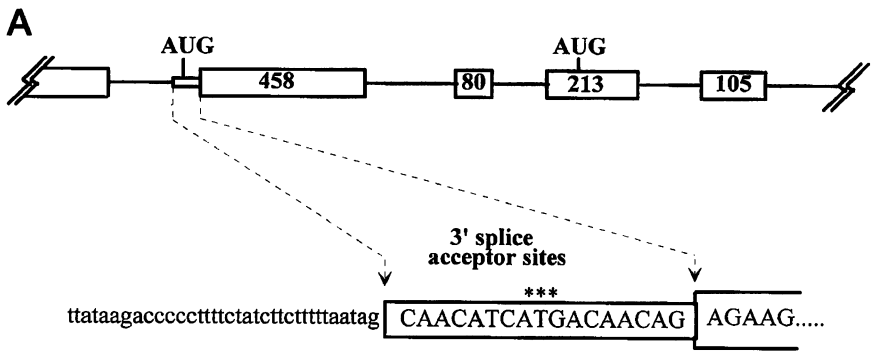


FIG. 5. Alternatively Spliced Exons Within the SAB. The diagram shows an enlargement of the SAB region of Figure 4. Exon 14 = 57 bp; Exon 15 = 42 bp; Exon 16 = 63 bp. See text for details. Only the three prominent isoforms (+++), (—), and (–+) are shown.

exon 16 (–+) accumulates in abundance (26). Indeed, late erythroblasts and reticulocytes contain only mRNA with the (–+) pattern of exon retention.

Considerable effort has been focused on characterizing the functional implication of the highly tissue- and differentiation stage-specific splicing event that leads to selective expression of exon 16 in the major erythroid isoform. Protein 4.1 forms a ternary complex with erythrocyte spectrin and short actin filaments, and thereby amplifies and stabilizes the interaction of spectrin and F-actin (9–11). This stabilizes the latticework along the "horizontal" plane of the cytoskeleton-membrane subsurface. The presence of exon 16 in the 10 KD SAB domain and the fact that expression of this exon is induced in the late stages of erythroid differentiation led us to define more precisely the structural features of erythrocyte protein 4.1 that are required for formation of the spectrin/actin ternary complex (27).

We generated recombinant isoforms of protein 4.1 and SAB peptides that contained or lacked exon 16, and examined their ability to complex with erythrocyte spectrin and F-actin. A complete discussion of



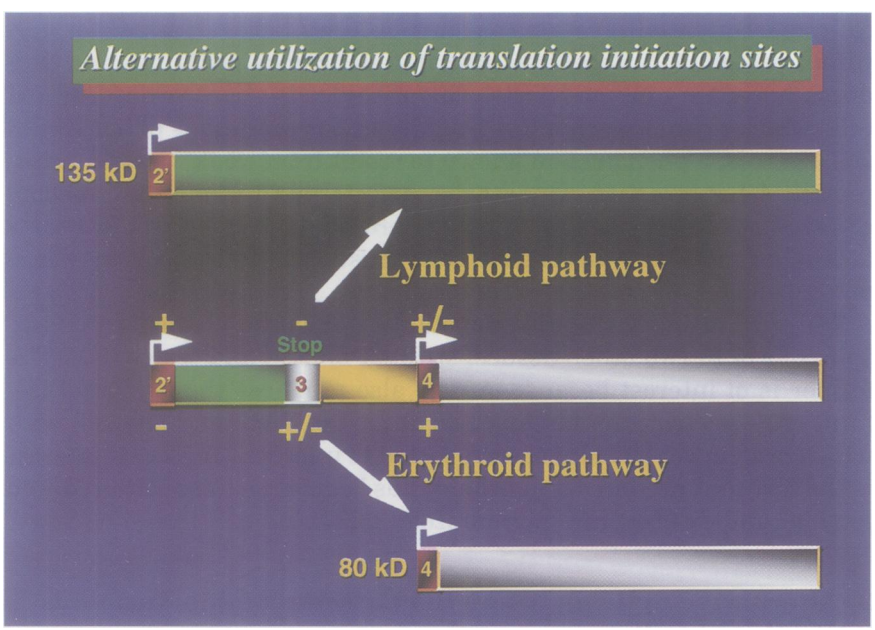


FIG. 6. Translation Patterns of Protein 4.1 mRNA. A. Nucleotide sequence of alternative 3′-splice acceptor sites. Note: The selection of the cryptic, upstream 3′ splice results in the inclusion of a 17 bp sequences containing the upstream translation initiation site. B. The diagram indicates how inclusion of the 17 bp motif (exon 2′) at the extreme 5′ end of the mRNA, and exclusion of the 80 bp exon (exon 3) containing stop codons allows for translation of a 135 KD form.

the experiments is beyond the scope of this paper (cf. Ref 27 for details), but the conclusion was clear: isoforms of either the complete protein or the 10 KD peptide domain alone bound spectrin-actin if they contained the 21 amino acid sequence encoded by exon 16. Isoforms

that lacked that sequence, but were otherwise identical, did not. These results suggest that interaction with spectrin- and F-actin is a specialized property of the Protein 4.1 isoform that is generated by the regulated splicing pathway activated during differentiation of the erythrocyte, a conclusion that has been confirmed by another group (28). This regulated splicing mechanism is crucial for the biogenesis of the highly specialized membrane that allows red cells to survive the mechanical stresses of the circulation.

We have also recently shown (29) that the *absence* of exon 16 precludes binding to *either* erythroid or nonerythroid spectrin. The predominant (—) forms of SAB expressed in nonerythroid tissues must then have functions other than spectrin-actin binding. This is highly important, since spectrin-actin binding is the only pathophysiologically consequential function that has been well characterized for erythroid protein 4.1. Complete inherited deficiency of the protein produces hereditary elliptocytosis. This disorder is characterized by red cells which exhibit mechanical fragility under conditions of horizontal shear stress, a phenomenon known to result from weak spectrin-actin-protein 4.1 interaction. Qualitative defects due to inherited splicing errors that abolish exon 16 inclusion produce the same result. Forms lacking exon 16 in nonerythroid cells must then have novel functions not dependent on spectrin-actin binding.

We have recently detected substantial amounts of the "erythroid" spliceoform in two other tissues: striated muscle and testes (30). The former is of interest because muscle also produces the erythroid isoform of β-spectrin, the binding target for the SAB, and myosin, with which protein 4.1 has also been reported to interact (31,32). Of interest is the fact that protein 4.1 deficiency has been associated with reduced fertility and dysmorphic spermatozoa (33,34).

2) Alternative Splicing Alters the Extreme Amino Terminus in Erythroid Cells

The prototypical erythroid form of protein 4.1 mRNA has an open translation-reading frame extending well into the 5′ untranslated region (25). The region is not translatable because it is interrupted by an 80 base pair exon (exon 3) that contains stop codons in all three reading frames. Thus, in erythroid cells, an 80 KD protein initiated from an initiator methionine codon in exon 3 is synthesized. We (19,23;35) and others (36) have shown that a 17 bp motif (exon 2′) is present in a subset of mRNAs in nonerythroid tissues. This motif contains a methionine codon surrounded by a strong consensus signal

for translation initiation. In these nonerythroid mRNAs, the 80 base pair sequence block (exon 3) containing the translation-stop codons is also invariably absent (spliced out). These two concerted mRNA splicing events (inclusion of exon 2′ and omission of exon 3) provide an alternative upstream translation start site with an open translation-reading frame in perfect register with the reading frame originating at the "erythroid" start site in exon 3. The resulting mRNA possesses two potentially usable translation start sites: the novel "upstream" AUG, and the "downstream" AUG utilized for synthesis of the 80 KD erythroid isoform. Use of the upstream site leads to a novel high molecular weight (135 KD) protein 4.1 isoform characterized by the presence of a 209 amino acid amino terminal extension, or "headpiece," attached to the 80 KD proteins (Figure 6).

The functions of the 135 KD forms are unknown. Importance is implied by their presence in many species, and by the fact that splicing events altering production of the headpiece are strictly regulated, at least during erythroid differentiation. Primitive erythroid cells contain the 135 KD mRNA and protein, but this splicing pathway is completely suppressed during later stages of maturation (19,23,36). Synthesis of protein 4.1 in red cells is thus completely dependent upon the downstream AUG codon.

We and others (37–39) have shown that mutations abolishing the downstream AUG cause selective deficiency of 80 KD protein 4.1 in red cells. Nonerythroid tissues from these patients contain only the 135 KD form. The patients are normal except for the red cell anomalies associated with HE, implying that: 1) the 135 KD form is important in nonerythroid tissues; 2) its absence is desirable in red cells; and 3) it is capable, in nonerythroid cells, of mimicking the function of the 80 KD isoforms. These patients also illustrate the point that mutations impairing a single copy gene expressed in multiple tissues can cause a clinical phenotype expressed in only one or a few tissues, provided that the mutation involves selective post-transcriptional events such as the selective use of translation start sites just noted.

Antibodies raised against the headpiece react with 135 KD proteins found both in the cytoplasm and the nucleus; additional peptides ranging in size from 30–160 KD also react (40,41). Cells arrested in mitosis show staining along the mitotic spindle fibers and the organizing centers, but not in chromatin (35,36). Our studies also indicate that the 135 KD isoform interacts with two nuclear proteins, elongation factor 1α (EF1α) and NuMA. EF1α is implicated in actin binding and bundling, microtubule severing, and association with the centrosome and mitotic apparatus (42). NuMA plays a role in the organization of

spindle microtubules (43). The 135 KD isoform may thus be involved in the organization of cell division and mitosis.

SUMMARY AND CONCLUSIONS

Regulated alternative pre-mRNA splicing is neither as widely appreciated as a fundamental aspect of controlled gene expression nor as thoroughly studied as transcriptional regulation. However, as exemplified by the phenomena cited in this review, alternative splicing is a fundamentally important mechanism used in the eukaryotic world to enhance the range, versatility and plasticity of the structural information contained within a gene, and to create additional strategies by which the net quantitative output of a given gene product can be controlled. Regulation of RNA splicing gives genes a modularity that adds flexibility, and, therefore, selective advantage, to eukaryotes. It is likely, though unproven, that this opportunity for refined regulation and diversification provides at least one basis for the existence of the tandem exon-intron-exon structure found in the vast majority of eukaryotic genes and many viral genes.

Many examples of alternative splicing are known, but, for the majority, no obvious biological impact of the alternatively spliced proteins on known cellular functions can be appreciated. Examples by which selectively regulated splicing pathways alter both the physiology and pathology of a major cellular event, such as differentiation and mechanical function of the red cell membrane, are thus relatively rare. The protein 4.1 gene and mRNA products thus provide an instructive and unusual system in which to explore the broader issue of the role of these regulatory mechanisms in the overall scheme of gene regulation and adaptation. The fact that hereditary hemolytic anemias result from mutations that directly or indirectly disrupt the splicing system emphasizes the relevance of these mechanisms to molecular medicine. The features of splicing that we have reviewed in this paper, and the specific impact that regulated splicing exerts on differentiating red cells have, we hope, convinced the reader that RNA splicing is an important, fascinating, and potentially fruitful area for future study of human disease processes.

BIBLIOGRAPHY

1. Benz, E. J., Jr., Principles of Molecular Biology, in Scientific Basis of Transfusion Medicine, Anderson, K.C., and Ness, P.M., W.B. Saunders, Philadelphia 1994;37–758.
2. Green, M.R. Biochemical mechanisms of constitutive and regulated pre-mRNA splic-

ing. Annu. Rev Cell Biol 1991;7:559–599.
3. Moore, MJ, Query CC, and Sharp PA. Splicing of pre-cursors to messenger RCAS by the spliceosome. In the RNA world (ed. R.F. Gesteland and J.F. Atkins), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1993:303–358..
4. Sharp PA. Split genes and RNA splicing. Cell 1994;77:805–815.
5. Rio DC. Splicing of pre-mRNA: Mechanism of regulation and role in development. Curr Opin Genet Dev 1993;3:574–584.
6. Kuo HC, Nassim FUH, and Grabowski PJ. Control of alternative splicing by the differential binding of U1 small nuclear ribonucleoprotein particle. Science 251: 1991;1045–1050.
7. Barker BS. Sex in flies: The splice of life. Nature 1989;340:521–524.
8. Mercer RW, Schneider JW, Benz EJ Jr: Na+,K+-ATPASE Structure: Red Blood Cell Membranes: Structure, Function, and Clinical Implications. Agre P and Parker J eds., Marcel Dekker, Inc., New York. 1989, pp 135–165.
9. Benz EJ Jr. The erythrocyte membrane and cytoskeleton: Structure, function, and disorders. In Stamatoyannopoulos G, Neinhuis AW, Majerus P, Varmus H, eds, The molecular basis of blood diseases, 2nd ed; 1994;257–292.
10. Conboy JG. Structure, function, and molecular genetics of erythroid membrane skeletal protein 4.1 in normal and abnormal red blood cells. Seminars in Hematology 1993;*30*(1):58–73.
11. Mohandas N, and Chasis JA. Red blood cell deformability, membrane material properties and shape: Regulation by transmembrane, skeletal and cytosolic proteins and lipids. Seminars in Hematology 1993;*30*(3):171–192.
12. Palek J, and Jarolim P. Red cell membrane disorders. In Hoffman R, Benz EJ, Shattil SJ, Furie B, Cohen HJ, Silberstein LE, eds, Hematology: Basic principles and practice, 2nd ed; 1995;667–709.
13. Leto TL, and Marchesi VT. A structural model of human erythrocyte protein 4.1. J Biol Chem 1984;*259*:4603–4608.
14. Correas I, Leto TM, Speicher DW, and Marchesi VT. Identification of the functional site of erythrocyte protein 4.1 involved in spectrin-actin associations. J Biol Chem 1986;*261*:3310–3315.
15. Anderson RA, Correas I, Massucco C, Castle JD, and Marchesi VT. Tissue-specific analogues of erythrocyte protein 4.1 retain functional domains. J Cell Biochem 1988;*37*:269–284.
16. Cohen CM, Foley SF, and Korsgren C. A protein immunologically related to erythrocyte band 4.1 is found on stress fibres of non-erythroid cells. Nature 1982;*299*:648–650.
17. Granger BL, and Lazarides E. Membrane skeletal protein 4.1 of avian erythrocytes is composed of multiple variants that exhibit tissue-specific expression. Cell 1988;*53*:601–615.
18. Leto TL, Pratt BM, and Madri JA. Mechanisms of cytoskeletal regulation: Modulation of aortic endothelial cell protein band 4.1 by the extracellular matrix. J Cell Physiol 1986;*127*:423–431.
19. Tang TK, Leto T, Marchesi VT, and Benz EJ Jr. Heterogeneity of mRNA and protein products arising from the protein 4.1 gene in erythroid and nonerythroid tissues. J Cell Biol 1990;*110*:617–624.
20. Ngai J, Stack JH, Moon RT, and Lazarides E. Regulated expression of multiple chicken erythroid membrane skeletal protein 4.1 variants is governed by differential RNA processing and translational control. Proc Natl Acad Sci USA 1987;*84*:4432–4436.
21. Conboy JG, Chan J, Mohandas N, and Kan YW. Multiple protein 4.1 isoforms

produced by alterantive splicing in human erythroid cells. Proc Natl Acad Sci USA 1988;*85*:9062–9065.

22. Tang TK, Leto TL, Correas I, Alonso MA, Marchesi VT, and Benz, EJ, Jr. Selective expression of an erythroid-specific isoform of protein 4.1. Proc Natl Acad Sci USA 1988;*85*:3713–3717.
23. Tang TK, Mazzucco CE, Leto TL, Marchesi VT, and Benz EZ, Jr. Translation of an mRNA species encoding nonerythroid protein 4.1 generates two proteins, one of which may be localized in the nucleus. Clin Res 1988;*36*:A405.
24. Tang CJ, and Tang TK. Rapid localization of membrane skeletal protein 4.1 (EL1) to human chromosome 1p33–p34.2 by nonradioactive in situ hybridization. Cytogenet Cell Genet 1991;*57*:119.
25. Baklouti F, Huang SC, Vuillamy T, Delaunay J, and Benz EJ Jr. Organization of human protein 4.1 genomic locus. New insights into the tissue-specific alternative splicing of the pre-mRNA. Genomics (In press). 1996.
26. Baklouti F, Huang SC, Tang TK, Marchesi VT, and Benz EJ Jr. Asynchronous regulation of splicing events within protein 4.1 pre-mRNA during erythroid differentiation. Blood 1996;*87*:3934–3941.
27. Horne W, Huang SC, Becker PS, and Benz EJ Jr. Tissue-specific alternative splicing of protein 4.1 inserts, an exon necessary for the formation of the ternary complex with erythrocyte spectrin and F-actin binding. Blood 1993;*82*:2558–2563.
28. Discher D, Patra M, Conboy JG, and Mohandas N. Mechanochemistry of the alternatively spliced spectrin-actin binding domain in membrane skeletal protein 4.1. J Biol Chem 1993;*268*(10):7186–95.
29. Huang SC, Frye CS, and Benz EJ Jr. Identification of the functional sites of nonerythroid protein 4.1 involved in fodrin-actin associations. Submitted to J Cell Biol., 1996.
30. Huang JP, Tang CJ, Kou GH, Marchesi VT, Benz EJ Jr, and Tang TK. Genomic structure of the locus encoding protein 4.1. Structural basis for complex combinational patterns for tissue-specific alterantive RNA splicing. J Biol Chem 1993;268(5):3758–66.
31. Racusen RH. Microscale, filtration-type binding assay for studying myosin-erythrocyte protein 4.1 interactions. Analytical Biochemistry 1990;*188*(2):344–348.
32. Pasternack GR, and Racusen RH. Erythrocyte protein 4.1 binds and regulates myosin. Proc Natl Acad Sci USA 1989;*86*(24):9712–9716.
33. Rousseaux-Prevost R, Lesur P, Collier F, Rigot JM, and Dalla Venezia N. Abnormal expressionof protein 4.1 in spermatozoa of infertile men with teratospermia. Lancet 1994;*343*(8900):764–765.
34. DeCesaris P, Filippini A, Stefanini M, and Ziparo E. Spectrin, fodrin and protein 4.1-like proteins in differentiation rat germ cells. Differentiation 1989;*41*(3):216–222.
35. Marchesi VT, Huang S, Tang TK, and Benz EJ Jr. Intranuclear Localization of high molecular weight isoform of protein 4.1 generated by alternative mRNA splicing. Blood 1990;*76*:12A.
36. Chasis JA, Coulombel L, Conboy J, McGee S, Andrews K, Kan YW, and Mohandas N. Differentiation-associated switches in protein 4.1 expression. Synthesis of multiple structural isoforms during normal human erythropoiesis. J Clin Invest 1993;*91*(1):329–38.
37. Dalla Venesia N, Gilsanz F, Alloisio N, Ducluzeau MT, Benz EJ Jr, and Delaunay J. Homozygous 4.1(-) hereditary elliptocytosis associated with a point mutation in the downstream initiation coon of protein 4.1 gene. J Clin Invest 1992;*90*(5):1713–1717.
38. Conboy JG, Mohandas N, Tchernia, and Kan YW. Molecular basis of hereditary

elliptocytosis due to protein 4.1 deficiency. N Engl J Med 1986;*315*:680–685.
39. Conboy JG, Shitamoto R, and Parra M. Hereditary elliptocytosis due to both qualitative and quantitative defects in membrane skeletal protein 4.1. Blood 1991;*78*:2438–43.
40. Correas I. Characterization of isoforms of protein 4.1 present in the nucleus. Biochem J 1991;*279*, 581–585.
41. Malek SN, Katumuluwa AI, and Pasternack GR. Identification and preliminary characterization of two related proliferation-associated nuclear phosphoproteins. J Biol Chem 1990;*265*(22):13400–9.
42. Marchesi VT, and Ngo N. In vitro assembly of multiprotein complexes containing α,β,and γ tubulin, heat shock protein HSP70, and elongation factor 1α. Proc Natl Acad Sci USA 1993;90:3028–3032.
43. Compton DA, and Cleveland DW. NuMA is required for the proper completion of mitosis. J Cell Bio 1993;120:947–957.