Brain and Muscle Express a Unique Alternative Transcript of αII Spectrin†

Carol D. Cianci, Zhushan Zhang, Deepti Pradhan, and Jon S. Morrow*

Department of Pathology, and the Department of Molecular, Cellular and Developmental Biology, Yale University, 310 Cedar Street, New Haven, Connecticut 06510

Received June 24, 1999; Revised Manuscript Received September 20, 1999

ABSTRACT: Alternative splicing of pre-mRNA transcripts of α and *β* spectrin has emerged as an important generator of diversity in this gene family, yet the functional consequences and extent of this diversity remains unknown. We have cloned and characterized full-length αII spectrin cDNA from human fetal brain (GenBank U83867 and U26396). On the basis of the predicted amino acid sequence, 11 amino acid substitutions, presumably representing polymorphisms, have been identified that distinguish this αII spectrin from human lung fibroblast αII spectrin. In addition, human fetal brain spectrin displays a novel five amino acid insertion in repeat 15 that arises from alternative mRNA splicing and that distinguishes this spectrin from lung fibroblast αII spectrin. This discovery, together with two previously identified regions of alternative mRNA splicing in αII spectrin suggest that as many as eight different splice forms of the mature protein might exist if all combinations (at inserts 1, 2, and 3) of alternative mRNA splicing are utilized. To assess this possibility, the tissue distribution of alternative exon usage was investigated by semiquantitative PCR with intron-jumping primer sets. Tissues examined were from mouse and included heart, kidney, lung, liver, thymus, spleen, brain, ovary, testis, and skeletal muscle, as well as mouse embryonic tissue. Transcripts both with and without insert 1, representing a 60 bp insertion within αII spectrin repeat 10, were identified in all tissues. In contrast, transcripts with insert 2, the novel 15 bp insertion reported here, were only expressed in brain, heart, skeletal muscle, and embryonic tissue. In all tissues examined only transcripts positive for insert 3, an 18 bp insertion in repeat 21, were amplified, even under conditions in which a 30% level of insert 3 negative transcript could be easily detected in artificially prepared control samples. All combinations of insert 1 and insert 2 were identified together in individual transcripts, verifying at least four distinct isoforms of αII spectrin. These have been named αIIΣ1 through αIIΣ4, in accord with current spectrin naming conventions. Dynamic molecular modeling of the 15th repeat unit incorporating insert 2 predicts that the spliced sequence forms a loop between helices A and B, and suggests that this insert might constitute a novel protein interaction site. The presence of this sequence in αIIΣ3 and αIIΣ4 spectrin suggests a specialized and heretofore unanticipated function for the 15th repeat of this molecule.

The spectrin-actin cytoskeleton, first identified in the erythrocyte, is a ubiquitous and central component of the machinery by which cells organize their surface and internal membranes (for reviews, see refs *1*−*4*). As nonerythroid cells have been examined, an increasing appreciation of the molecular diversity of the spectrin skeleton has emerged. In the human erythrocyte, the spectrin functional unit is composed of two nonidentical but homologous subunits, termed αI and *β*I, arising from genes on chromosomes 1 (*5*) and 14 (*6*), respectively. At the plasma membrane of most nonerythroid tissues is a similar spectrin heterocomplex, composed of αII and *β*II subunits, the product of genes on chromosomes 9 (*7*) and 2 (*8*), respectively. Associated with the Golgi and other vesicular compartments of the cell is *β*III spectrin, a third form of unique spectrin that arises from a gene on chromosome 11 (*9*). Additional forms of spectrin also may exist, since antigenically unique isoforms have been detected in association with the acetylcholine receptor in muscle cells (*10*, *11*), and spectrins other than *β*III may also be present in the Golgi compartment of epithelial cells (*12*, *13*). Beyond distinct gene utilization, additional spectrin diversity arises from alternative mRNA splicing. Recognized human spectrin isoforms include two forms of *β*I spectrin, derived from differential exon utilization at the COOH-terminus (*14*−*18*); two *β*II isoforms, derived from alternative exon utilization at the NH_2-terminus (*19*; Forget et al., unpublished observations); and two regions of alternative splicing in αII spectrin, one in the tenth repeat unit and the other in repeat unit 21 (*15*). With the possible exception of COOH-terminal alternative splicing in *β*I spectrin, which places a pleckstrin homology domain in *β*IΣ2 spectrin that is capable of binding $PtdInsP_2$ (*20*, *21*), the functional consequences of this structural diversity remains unclear.

The present report describes the complete cloning and characterization of human fetal brain αII spectrin cDNA. In addition to several single-base divergences from human lung fibroblasts spectrin αII cDNA sequence, that presumably represent polymorphism's in this gene, we identify a novel 15 bp region introduced by alternative mRNA splicing after codon 1586 (insert 2). This insertion is predicted to

* To whom correspondence should be addressed. Phone: (203) 785-3624. Fax: (203) 785-7037. E-mail: jon.morrow@yale.edu.

† This work was supported by grants from the National Institutes of Health to J.S.M.

10.1021/bi991458k CCC: $18.00
Published on Web 11/05/1999

encode a five residue insertion in αII-spectrin's repeat unit 15. Dynamic molecular modeling of this insertion predicts that the novel (inserted) sequence extends the interhelix loop between helices A and B of the 15th repetitive unit. The tissue utilization of insert 2, as well as the utilization of two other previously described inserts that also arise by alternative mRNA processing, are evaluated by RT-PCR. Surprisingly, we find that insert 2 generates a unique isoform of αII spectrin expressed only in brain, muscle, and embryonic tissue, while the other transcripts are more widely distributed. These results thus identify a signature sequence that is unique to brain and muscle αII spectrin. We hypothesize that this sequence bestows on αII spectrin a novel ligand-binding capacity.

The cDNA sequence of human fetal brain spectrin has been deposited in GenBank as accession numbers U83867 and U26396.

MATERIALS AND METHODS

All DNA manipulations, including PCR amplifications, restriction enzyme digestion's, plasmid preparation, vector construction, subcloning and sequencing, followed standard methods unless otherwise specified (*22*).

Cloning and Characterization of Human Fetal Brain αII Spectrin. The isolation of cDNA clones encoding αII spectrin from a human fetal brain library (Stratagene) in λ-ZAP has been previously reported (*23*). The initial probe used to identify fetal spectrin clones were isolated by hybridization with an α-[^{32}P]dCTP random labeled 1.7 Kb PCR fragment constructed from primers based on the published human lung αII spectrin sequence (nucleotides 2455−4211) (*15*). All positive purified plaques were subcloned in vivo into pBluescript directly from the λ ZAP vector using helper phage R408, following the protocol accompanying the λ ZAP product literature. Clone sizes were approximated by restriction digests with *Eco*R1. The ends of each clone were sequenced using T7 and T3 primers, and the overlapping clones were aligned. Clones used for assembly were expanded, and large-scale plasmids were prepared (Qiagen) and completely sequenced by the dideoxy chain-termination methods using Sequenase v 2.0 (United States Biochemical) (*24*). Primer walking and extension was used to complete the sequence of the full-length gene. At least two independent clones were sequenced in both directions to confirm regions containing polymorphic nucleotides.

Isoform Analysis. Total mouse RNA was obtained from Ambion (Austin, TX). Reverse transcription was performed on 5 μg of total RNA with either oligo-dt$_{(12-18)}$ or random hexamers, utilizing the Superscript II Second Strand Amplification System (Life Technologies). PCR was performed in a 100 μL reaction mixture using 5 μL of the RT-cDNA, 0.5 μg of the specific internal primers, 0.5 μL *Taq* polymerase (Perkin-Elmer), 1.5 mM $MgCl_2$, and 200 mM dNTP's. Primers were chosen so as to flank known regions of alternative splicing, and hence were "intron-jumping". The specific primers used for each transcript were insert 1, forward primer (5′-GGTGGAAAGTGGAAGTGAACGATC-3′), reverse primer (5′-TGCTGTAGTTCATTCGCTTCACGG-3′); insert 2, forward primer (5′-GGATGAGATTGAGGCT-TGGATC-3′), reverse primer (5′-GTCGATAACCCCACG-GATCCGG-3′); insert 3, forward primer 5′-GGAGTTTGC-CCAGCACGCCAACG-3′, reverse primer (5′-GCTGACT-TCTCATGGCTCGGA−3′).

PCR reactions were performed in a Perkin-Elmer thermocycler using a "Touchdown Low" program, during which the annealing temperature was reduced from 50 to 42 °C in increments of 2 °C/cycle, followed by 30 cycles at a 42 °C annealing temperature. The PCR products were analyzed using 12% Tris-borate EDTA (TBE)[1] polyacrylamide gels and visualized by ethidium bromide and ultraviolet light. All PCR products were immediately subcloned into the pCR2.1 (TA) cloning vector from Invitrogen (San Diego, CA), for subsequent sequencing to verify the validity of the amplified sequences.

PCR Calibration. To quantitatively assess the levels of transcript present in the various tissues for each spectrin isoform, control mixtures containing known quantities of cDNA, representing the presence or absence of each insert, were prepared and evaluated over a range of PCR amplification and gel loading conditions. Favorable conditions were identified (as above) that yielded an approximately linear amplification response for all αII inserts and which were within a linear range of densitometric responsiveness after ethidium bromide staining of the gels. Using these conditions, a series of 10 standard samples, containing mixtures of insert positive vs insert negative cDNA in ratios ranging from 0:1 to 1:0, were evaluated in conjunction with each tissue insert analysis, and the standard curves derived from these amplifications were evaluated by nonlinear regression analysis. The resultant fitting coefficients were then used in the evaluation of the isoform ratios in the tissues.

Analysis of Spectrin Isoform Pairing. Mouse embryonic mRNA was RT-PCR'd by oligo(dt) using Superscript Second Strand Amplification System (Life Technologies). The various primers used are summarized in Table 1. A PCR reaction was initially performed using upstream primers which encompassed or deleted the insert 1 sequence (primers number 34 105 and 34 106, respectively) and a downstream primer (primer number 13 623), which encompassed insert 2. These PCR reactions were then analyzed by southern blotting using sequence specific oligos as probes. The oligos (all approximately 50 bp in length) were labeled by Amersham's direct nucleic acid labeling and detection system (RPN) according to manufacture's protocol. To maintain the specificity and the minimum length (50 bp) of the probes for this method of labeling, irrelevant sequence was added to the ends of certain primers (indicated in by underline). Oligo's used as probes which were in the minus (−) form were constructed by looping out the insert sequence.

Molecular Modeling. The predicted structural consequences of the alternative sequences in αII spectrin were modeled following the procedure used to predict the structure of the calpain cleavage site in αII spectrin (*23*). Briefly, the sequence of a repeat of interest was aligned by the program Best fit (*25*) with the sequence of the 14th repeat unit of *Drosophila* αII spectrin, for which the three-dimensional structure has been determined (*26*). This allowed the assignment of specific residues to each of the three helices in the spectrin repeat unit. These residues were then substituted at their homologous positions in the three-dimensional crystal

[1] Abbreviations: RPN, Amersham's direct nucleic acid labeling and detection system; TBE, 12% Tris-borate EDTA.

Table 1: Oligonucleotides Used for RT-PCR Analysis of Specific Spectrin Isoforms

name	identity	nt	sequence (5′-3′)	use
41 846	insert 2 + forward	4742	ATCCCACCAACATCCAGCTTTCCAAGCTGCTGAGCAAGCACCAGAAGCAC	probe
41 947	insert 2 + reverse	4791	GTGCTTCTGGTGCTTGCTCAGCAGCTTGGAAAGCTGGATGTTGGTGGGAT	probe
41 848	insert 1 − forward	3140	ATGCATGCAGCAGATTGACAATTCAGTATCATTCTCTGCTGGAATGCATGC	probe
41 849	insert 1 − reverse	3234	GCATGCATTCCAGCAGAGAATGATACTGCTTGTCAATCTGCTGCATGCAT	probe
41 850	insert 2 − forward	4742	GATCGATCATCCCACCAACATCCAGAGCAAGCACCAGAAGCAGATCGATC	probe
41 851	insert 2 − reverse	4789	GATCGATCGCTTCTTGGTGCTTGCTCTGGATGCTTGGTGGGATGATCGATC	probe
41 665	insert 1 + forward	3157	ACACGCATAACTAAGGAGGCCGGCAGTGTATCTCTGCGTATGAAGCAGGTGGAAGAACTA	probe
41 670	insert 1 + reverse	3204	TAGTTCTTCCACCTGCTTCATACGCAGAGATACACTGCCGGCCTCCTTAGTTATGCGTGT	probe
34 105	insert 1+ forward	3148	GACAATCAGACACGCATAAC	PCR
34 106	insert1 − forward	3148	GACAATCAGTATCATTCTCTGCTG	PCR
13 623	αII downstream reverse	4925	CCAGGCGGGCCTTGACAGCATCCTC	PCR

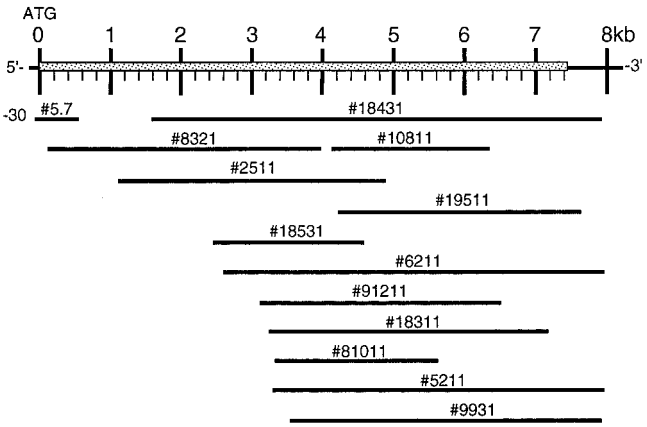

FIGURE 1: The cloning of human fetal brain αII spectrin. Screening of a human fetal brain library for αII spectrin identified 39 clones. Representative clones used to establish and verify the complete cDNA sequence of the brain isoform of αII spectrin are depicted. The complete cDNA sequence derived from the clones shown may be accessioned from GenBank, U83867.

structure of the *Drosophila* spectrin; also inserted but not structured were the additional residues provided by the alternative splice. This entire structure was then placed in a 6 Å water shell and dynamically modeled and energy minimized as before (for more complete description of the method used, see refs *23* and *27*). Resultant structures were displayed using the Biosym software package operating on a Silicon Graphics Indigo^2 Extreme workstation.

RESULTS

Cloning and Characterization of Human Fetal Brain αII *Spectrin.* Previously, we have identified a series of 39 overlapping clones encoding human fetal brain αII spectrin, and have characterized one of these (clone 18 531) as encoding the regions of αII spectrin involved with calmodulin binding and μ-calpain cleavage (*23*). In the present study, we extend our analysis of these clones and present the complete structure of this gene. Thirty-nine positive clones were obtained spanning the entire length of the published αII human lung sequence, except for the N-terminal first 75 nucleotides (Figure 1). The 5′-end of the gene was obtained by PCR using an upstream primer derived from the 5′-untranslated region of the lung sequence and a reverse sense primer starting at nt 547, encompassing the unique *Mun*I site, thereby allowing the convenient assembly of the full-length αII product. The entire full-length cDNA was sequenced from the assembled product (Figure 2). The derived amino acid sequence of human brain αII spectrin contained 2477 residues with a predicted molecular mass of 285 095 Da. This is to be compared with human lung fibroblast sequence with a predicted 2472 residues and a calculated molecular mass of 283 964 Da; chicken brain αII spectrin sequence of 2477 codons with a molecular mass of 285 369 Da (*28*); and *Drosophila* αII-spectrin containing 2415 residues with a predicted mass of 278 364 Da (*29*).

Compared to the human lung sequence, human fetal brain spectrin displayed 19 nucleotide substitutions, several of which resulted in alternative amino acid usage (Figure 2 and Table 2). All nucleotide differences were confirmed in at least two and usually three independent clones from the fetal brain library. Several of these nucleotide substitutions also generated altered enzyme restriction sites, and were additionally verified in some cases by their sensitivity (or lack of sensitivity) to the appropriate enzyme (data not shown). Since it was also possible that the variances from the fibroblast sequence arose from sequencing errors in the latter, an effort was made to verify the sequence of the original fibroblast clones. Repeat sequencing of the only available original clone (clone 2.7A, kindly provided by Dr. Bernard Forget), representing bases 2128−4885, found no differences from the reported sequence. Thus, the divergences at least within this region (at nucleotides 2209, 2631, and 3899) must represent true polymorphisms. It cannot be definitely determined at this point whether the other divergences from the published lung fibroblast sequence represent true polymorphisms or sequencing errors in the original report.

When the sequence of αII spectrin is aligned with respect to the phasing of the structural repeat unit (*26*), several interesting features emerge (Figures 2 and 3). Putative sequence repeat units 9 and 10 together appear to form only a single structural repeat unit, with the SH3 domain and its immediate downstream flanking sequence arising within a segment joining helices B and C. Also within this structural unit is an alternative transcript (insert 1) that is predicted to arise by interrupting helix C in a fashion analogous to the positioning of the calpain cleavage and calmodulin binding site found in repeat 11 (*23*).

Human Fetal Brain αII *Spectrin Contains a Novel Inserted Sequence in Repeat 15.* Sequence analysis of the complete full-length αII spectrin gene detected an insertion of 15 nucleotides in repeat 15, resulting in a predicted amino acid sequence that contained an additional five residues. This sequence was not present in the reported human αII lung spectrin sequence nor in the sequence from *Drosophila*. Surprisingly, a comparison with the sequence of chicken brain spectrin revealed a homologous region inserted at

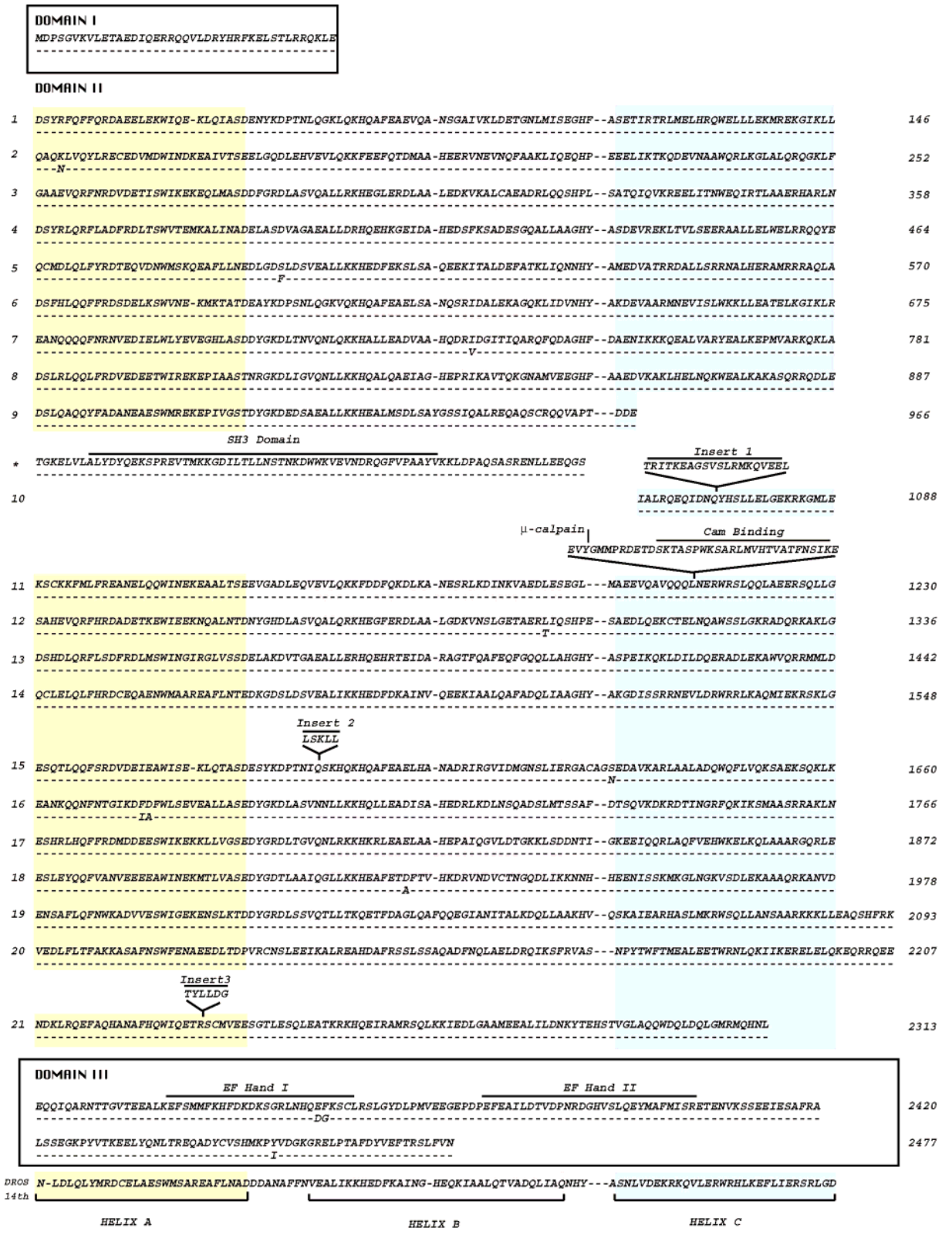

FIGURE 2: The derived amino acid sequence of human brain αII spectrin, compared to that of human lung αII spectrin (*15*). Residues (total = 11) of the lung isoform (dashes) that differ from those in fetal brain are shown. In this alignment, the sequence of human brain αII spectrin has been aligned so as to reflect the phasing of the triple helical repeat structure of each spectrin unit, based on the positioning of helices A (shaded gold), B (shaded red), and C (shaded blue) of *Drosophila* αII spectrin's 14th repeat unit [shown at the bottom (*26*)]. The number of each structural repeat unit is shown at the left. Residue positions are shown on the right. When aligned on the basis of structural repeats, it is apparent that repeat 9 consists of only 2/3 of a repeat unit, with the balance of this structural unit contributed by sequence in the terminal portions of putative repeat unit 10. Most of the other sequence in what was previously considered part of repeat 10 [(marked with an asterisk (*)] is nonhomologous, consisting of the SH3 domain and in some isoforms insert 1. Also of note is the positioning of insert 2 that arises near the junction between helices A and B in repeat 15, and insert 3 predicted to arise from helix A of repeat 21. The primary site cleaved by μ-calpain and the calmodulin binding domain (CaM) are indicated.

Table 2: Divergence between Human Brain and Lung αII Spectrin

nt[a]	ΔL−B	codon	amino acid
433	T−C	TTG-**C**TG	L to L
450	C−G	AAC-AA**G**	N to K
1493	T−C	TTC-T**C**C	F to S
1647	C−T	GAC-GA**T**	D to D
1767	G−C	GTG-GT**C**	V to V
2209	G−A	GTC-**A**TC	V to I
2631	A−C	GCA−GC**C**	A to A
3899	C−T	ACC-A**T**C	T to I
4889 (4874)	A−G	AAT-A**G**T	N to S
4938 (4923)	A−G	CAA−CA**G**	Q to Q
5023 (5008)	A−T	ATT-**T**TT	I to F
5027 (5012)	C−A	GCA−G**A**C	A to D
5028 (5013)	A−C		"
5768 (5753)	C−A	GCC-G**A**C	A to D
7056 (7041)	T−G	GAT-GA**G**	D to E
7057 (7042)	G−T	GGC-**TT**C	G to F
7058 (7043)	G−T		"
7357 (7342)	A−T	ATC-**TA**C	I to Y
7358 (7343)	T−A		"

[a] Nucleotide positions are numbered according to the brain αII spectrin sequence (lung positions in brackets). Abbreviations: (L), lung; (B), brain; Δ, change.

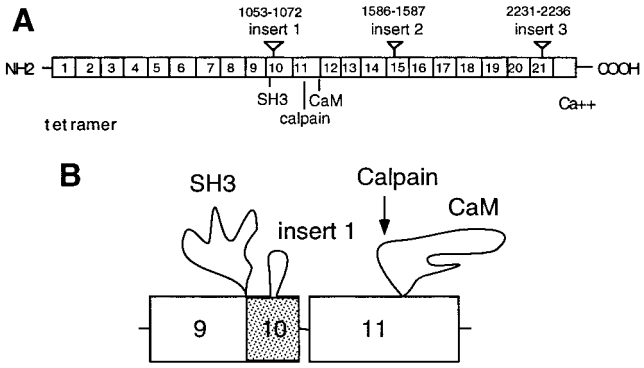

FIGURE 3: Three regions of alternative splicing identified in αII spectrin with their relationship to functional domains (cartoon) (A)-Schematic representation of the three regions of alternative splicing in human αII spectrin, and their relationship to the overall domain structure of spectrin. (B) Detailed schematic of the putative relationship of the nonhomologous sequences within repeats 9−11 to the alternatively spliced sequences in this region (insert 1) and the calmodulin-binding domain. In this region, it is proposed that a single basic spectrin repeat unit is formed from homologous sequences in repeats 9 and 10, and this unit is joined to repeat 11 which displays a nonhomologous calmodulin binding and calpain cleavage site within a nonhomologous sequence proposed to arise within helix C of repeat 11 (*23*).

precisely the same locus (*28*). Since both the chicken and human mRNAs with this sequence were derived from brain libraries, it seemed likely that this novel sequence might represent a relatively brain-specific transcript. To explore this hypotheses, as well as evaluate the tissue distribution of the other recognized splice forms of αII spectrin (Figures 2 and 3), all clones overlapping in the regions of repeat 11, repeat 15, and repeat 21, the sites of recognized alternative splicing events, were examined by cDNA sequencing for the presence or absence of alternative transcripts, and a variety of tissues were examined by RT-PCR for transcripts of these inserts (see below). A total of 10 clones encompassing repeat 15 were examined; in no instances were clones identified in the brain library that lacked inserts 1, 2, or 3, suggesting that the dominant αII transcript in brain contains all three regions of inserted sequence. A search of GenBank for other proteins that might display the 5 residue insert 2 motif revealed no homologies.

Insert 2 of αII *Spectrin Is Restricted to Brain, Muscle, and Embryonic Tissues.* Utilizing RT-PCR with primers flanking the regions of interest, total RNA from a series of mouse tissues was screened for the expression of the three known alternative transcript sequences (Figure 4). To calibrate these RT-PCR studies, a series of cDNAs either with or without insert 1, insert 2, or insert 3 were mixed in varying proportions, and then amplified and quantified by densitometry in parallel with the determination of the tissue samples. Precautions were taken to ensure that the densitometric signals were in a linear range of response; at times, this necessitated evaluating gels with different levels of loading. The resulting standard curve of optical density for each amplimer vs its relative abundance was then fit by nonlinear regression and the coefficients obtained used to evaluate the proportion of insert positive vs insert negative transcripts in each tissue extract. Transcripts with or without insert 1 were present in approximately equal amounts in heart, kidney, ovary, testis, and embryo. Lung, liver, thymus, spleen, brain, and skeletal muscle expressed predominately the insert(+) isoform (Figure 4B). Sequence analysis confirmed the anticipated structure of these PCR products. A larger PCR product evident in Figure 4 (marked by *) was also sequenced, and was found not to contain spectrin sequence. Transcripts with insert 2, the novel sequence identified here in human αII spectrin, were the only ones present in brain, while in heart muscle, skeletal muscle, and embryo, both insert 2(+) and insert 2(−) forms were detected. The remaining tissues (lung, liver, thymus, spleen, kidney, ovary, and testis) exclusively expressed transcripts lacking insert 2. Finally, in all tissues examined only insert 3, positive transcripts were identified; control studies indicated that insert 3 negative transcripts would have been detected if they were present at 30% or greater levels in total spectrin mRNA. Quantitation of these PCR products suggested that different tissues characteristically express different combinations of αII spectrin isoforms (Figure 4B), presumably to tailor the functional properties of their spectrin skeleton to each tissue's unique requirements. These results also clearly indicate that all combinations of splice forms bridging inserts 1 and 2 must exist. For example, in heart tissue, nearly 70% of αII spectrin carries insert 2, but less than 30% of this spectrin contains insert 1 (Figure 4B, lane 1). Therefore, αII spectrin with insert 2 but not insert 1 must exist. Similarly, in brain, almost all αII spectrin carries both insert 1 and insert 2 (Figure 4B, lane 7). In most tissues, most of the spectrin carries insert 1 but not insert 2; and finally, in embryonic tissue, over half of the spectrin contains neither insert 1 nor insert 2.

To verify these deductions, additional RT-PCR reactions were carried out on mRNA extracted from mouse embryos, in which the upstream primer was directed against either insert 1 itself or the sequence of the same region that lacks insert 1. The downstream primer in each case was sequence beyond insert 2. The resulting amplimers were probed by Southern blot for the presence or absence of inserts 1 and 2 (Figure 5). Also examined in these blots as a control were two cDNA constructs representing the region about the first alternative splice, with or without insert 1. As is evident from the EtBr stained gel (Figure 5, top), products representing

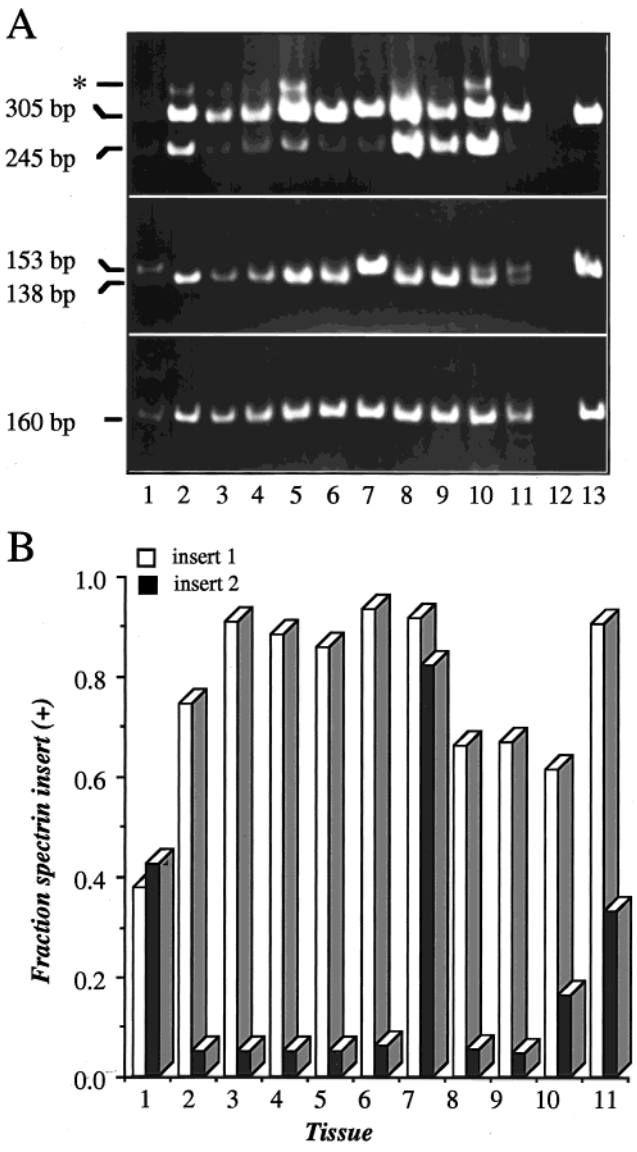

FIGURE 4: PCR analysis of αII transcripts in tissues (A) RT-PCR analysis of mouse tissue extracts, using primers specific for each of the three regions of alternative transcription in αII spectrin. Top panel, insert 1; middle panel, inset 2; bottom panel, insert 3. The mRNA for RT-PCR in each lane was derived from: lane 1, heart; 2, kidney; 3, lung; 4, liver; 5, thymus; 6, spleen; 7, brain; 8, ovary; 9, testis; 10, embryo; 11, skeletal muscle; 12, water only (negative control); 13, full-length αII spectrin cDNA (positive control). The size of the expected products for insert (+) or (−) transcripts at each locus is shown. The validity of all amplified products was verified by direct cDNA sequencing. A representative gel is shown. (B) Densitometric analysis of the PCR amplification products allowed quantitation of the abundance of each insert in the mRNA extracted from each tissue. In multiple analyses, the abundance of transcript was compared with standard curves determined by the parallel amplification of artificial mixtures containing a known quantity of each transcript. Precautions were also taken by analyzing gels loaded with different amounts of amplified cDNA, so as to ensure that each determination fell within the linear range of densitometric responsiveness. For each tissue, the fraction of αII transcripts that contained either insert 1 (open bars) or insert 2 (shaded bars) is shown. The tissue designations correspond to those in panel A.

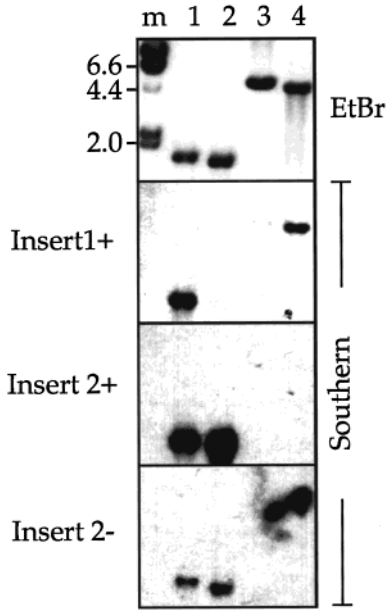

FIGURE 5: Inserts 1 and 2 appear independently in embryonic spectrin mRNA. The simultaneous presence of insert 1 and insert 2 in contiguous transcripts was analyzed by a two-step procedure in which primers (see Table 1) recognizing either insert 1 (lane 1) or the sequence created by its omission (lane 2) were used in conjunction with primers downstream of insert 2 to amplify by RT-PCR embryonic mouse mRNA. Southern blotting of the reaction products with oligonucleotides specific for insert 2+ or insert 2- cDNA were then used to determine what combinations of insert 1 and insert 2 appeared together. Also loaded on these gels were cDNA constructs representing residues 967−1077 of spectrin either with insert 1 (lane 4) or without insert 1 (lane 3). The EtBr-stained gel demonstrates that all cDNAs were loaded approximately equally. Note that regardless of whether the amplimer contains insert 1, the products hybridize with probes specific for both insert 2+ and insert 2− sequences. The irregular intensity near the position of the controls (lanes 3 and 4) for the blot with the insert 2 (−) probe are artifact; these constructs do not extend to the insert 2 region. The marker lane is "m".

both insert 1+ and insert 1− amplimers were generated in abundance, and each of these products hybridized with both insert 2+ and insert 2− specific oligonucleotides. These results thus directly demonstrate that all expected mRNA transcripts are present, representing either the presence or absence of insert 1 and insert 2 simultaneously in a single transcript. Thus, at least four specific isoforms of αII spectrin are transcribed.

Although no insert 3 negative transcripts were detected in these studies, a cDNA clone of such a transcript has been noted in one published study (*15*), and a nearly identical sequence constitutes exon 46 in the αI spectrin gene (*30*). Thus, although only identified in a single report, it remains likely that this transcript is real and represents an alternative splice when the corresponding exon (exon 46?) in αII spectrin is deleted. Our failure to detect transcripts without insert 3 suggests that deletion of insert 3 probably represents a rare message. The minimal detection limit for insert 3 in the calibrated PCR studies reported here was 30% of the total spectrin mRNA; while this limit sets a maximal level for insert 3 (−) transcripts, we suspect that in fact such transcripts in fact represent significantly less than 30% of total spectrin mRNA, based on our inability to detect such a transcript in any of the tissues examined even when using nested PCR amplification techniques (data not shown). If insert 3 (−) transcripts do appear with each of the four isoforms formed by combinations of insert 1 and insert 2, then in total there are eight isoforms of αII spectrin. We term these isoforms αII Σ1−8, as depicted in Table 3.

Table 3: Summary of αII Spectrin Isoforms

isoform	insert 1	insert 2	insert 3	calcd MW	tissue (most abundant)
αIIΣ1	yes	yes	yes	285 095	brain & muscle
αIIΣ2	yes	no	yes	284 539	most tissues
αIIΣ3	no	yes	yes	282 837	muscle & embryo
αIIΣ4	no	no	yes	282 282	gonads & embryo
αIIΣ5	yes	yes	no	284 431	not established
αIIΣ6	yes	no	no	283 876	not established
αIIΣ7	no	yes	no	282 174	not established
αIIΣ8	no	no	no	281 619	not established

Dynamic Molecular Modeling of the Repeat 15 Alternative Sequence. The fundamental structural feature of the conserved spectrin repeat unit is a triple α-helical motif (*26*). The novel insert described here arises in a region of conserved repetitive sequence within repeat 15, a region whose secondary and tertiary structural features presumably preserve this triple helical motif. It was therefore of interest to understand how the alternatively spliced sequence might be accommodated into the spectrin repeat structure. Primary sequence alignment of repeat unit 15 (with or without insert 2) with other αII spectrin repeats and with the sequence of the *Drosophila* αII spectrin 14th unit identified the portions of each repeat most likely to be involved with forming each of the three helices (Figure 1). From this analysis, two regions of sequence divergence in repeat 15 were apparent. The inserted sequence is predicted to lie within the region joining helix A to helix B. In addition, several divergent residues were noted within helix C of this repeat, at codon positions 1629−1630. Using the alignment depicted in Figure 1 and as a starting point for the modeling calculations, the sequence of human repeat 15 was superimposed onto the backbone three-dimensional structure of the *Drosophila* spectrin repeat. Energy minimization and dynamic molecular modeling algorithms were then used to estimate the probable structural consequences of each insertion, after first mathematically heating the subunit to 1000 °C to unfold all secondary and tertiary structure. In these computations, each repeat unit was modeled as if it were in a 6 Å water shell, a consideration found necessary in control studies to achieve a minimized structure by dynamic modeling that is faithful to the known structure of the *Drosophila* 14th repeat. This dynamic modeling approach has previously been used to estimate the structure of the calmodulin-binding and calpain-cleavage site in the 11th repeat unit of αII spectrin (*23*), as well as the structure of the spectrin self-association complex (*27*, *31*). The predicted structure derived from this approach is shown in Figure 6 for repeat 15 with and without the insert. The nonhomologous residues in helix C may distort helix C slightly, bringing its distal end into closer approximation to the region of alternative splicing, which is predicted to form a loop within the sequence joining helix A to helix B. The inserted sequence itself is predicted to cause a looping out of the segment joining helix A to helix B, so as to form an exposed turn with substantial solvent accessibility.

It is interesting to note that the structure predicted by this energy minimization procedure does not support an alternative possibility that insert 2 folds into an α-helical structure that extends either the A- or B-helix. It has been suggested that the two extra turns present in the B-helix (relative to the A or C helices) might stabilize the connection between two successive units (*26*). Given this, if insert 2 were to extend the A helix, it might complex with the B-helix and thereby reduce or altering intersubunit interactions. Multiple energy minimization analyses, beginning from different initial states and with reasonable variations in the amount of water in the hydration shell, all converged to essentially the same structure as depicted in Figure 6. Significantly, in no instance was the inserted sequence predicted to form or extend an α-helix. Inspection of this sequence reveals that it is highly hydrophilic, without the amphipathic character characteristic of spectrin's other helices. We therefore consider it unlikely that the modeling procedure is seriously misrepresenting the structure of this unit.

DISCUSSION

The complete cDNA sequence of an αII spectrin derived from human fetal brain has been determined, and 19 single-base alterations were identified that distinguish it from the only other characterized human αII spectrin gene (*15*). These polymorphisms create several novel restriction enzyme cleavage sites, as well as 11 amino acid substitutions in the predicted gene product. More importantly we have identified a novel 15 bp sequence inserted by alternative mRNA processing into αII spectrin's 15th repetitive unit. While the function of this sequence remains unknown, we postulate that it constitutes a unique binding for a still unknown ligand. This hypothesis is supported by preliminary surface plasmon resonance studies that appear to detect such a ligand in whole brain extracts (unpublished observations). That this insertion represents the product of alternative splicing is supported by (i) the presence of this insert in 10 independent clones isolated from the fetal brain spectrin library; (ii) the identity of the brain and lung αII spectrin cDNA sequences outside of the region of the insertion; (iii) the location of the inserted sequence precisely at a region homologous to the junction between exons 32 and 33 of the human αI spectrin gene, a gene thought to have arisen from αII spectrin by gene duplication (*14*); (iv) the presence of only a single αII spectrin gene, that appears on chromosome 9 (*7*); and (v) the presence of all combinations of insert 1 and insert 2 within single RT-PCR products that bridge both insertion sites. This novel insert appears to be exclusively expressed in brain, muscle, and embryonic tissues. Modeling of its three-dimensional structure suggests that it forms a binding site for an as of yet unidentified ligand, a result consonant with the preliminary findings cited above.

With this report, five α-spectrins have now been completely characterized: human fetal brain (this report); human lung fibroblast (*15*); human erythrocyte (*14*); chicken brain (*28*); and *Drosophila* (*29*). Collectively, these studies establish that the nonerythroid αII-spectrins are highly conserved, with at least 90% identity between vertebrate species and 63% identity between Drosophila and chicken. Conversely, human erythrocyte αI-spectrin shares only 50−60% identity with human αII-spectrin, presumably reflecting the highly specialized nature of mammalian erythrocytes, where it is most prominent. What is increasingly apparent from these studies is the diversity that arises from alternative mRNA splicing. Three sites of such splicing in αII spectrin have now been identified. In addition, as reported here, it appears that at least inserts 1 and 2 can each be utilized independently, creating at a minimum four different αII spectrin transcripts with tissue specific patterns of expression. In

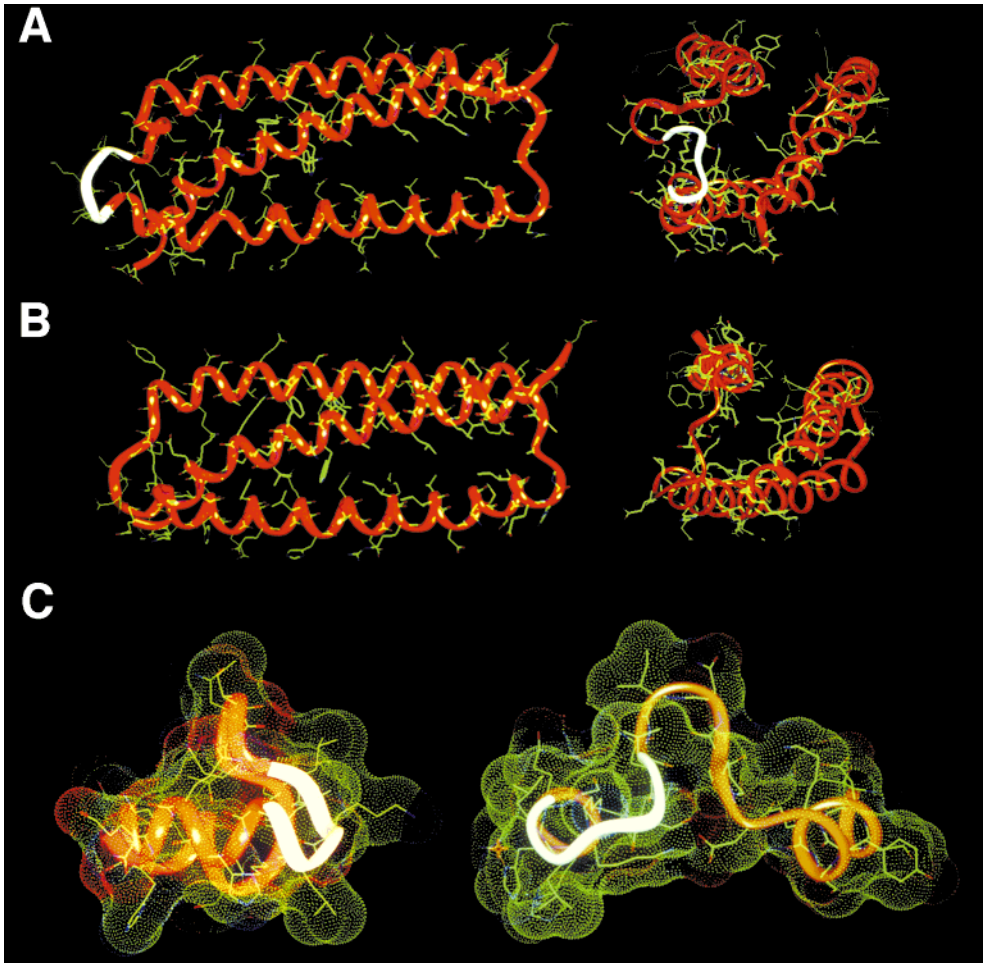

FIGURE 6: Dynamic molecular modeling places insert 2 between helices A and B of the spectrin structural repeat. The sequence of the 15th repeat unit of αII brain spectrin was aligned with the 14th repeat unit of *Drosophila* alpha spectrin to determine the probable boundaries of each helix in the repeat unit (see Figure 1). Thereafter, the structure of the human 15th repeat was approximated by dynamic molecular modeling and energy minimization, as described in Materials and Methods. **(A)** Two-axis views of the estimated structure of αII spectrin repeat 15 with insert 2 (white), positioned between helix A (at top of structure) and helix B (at bottom of structure). Helix C (diagonal, across structure) is predicted to bend relative to the *Drosophila* structure due to the presence of a nonhomologous insertion (see Figure 1). This bend appears to move the point of interaction of the distal end of helix C slightly toward helix B, exposing the region of linkage between helices A and B and the site of insert 2. **(B)** Two-axis views of the estimated structure of αII spectrin, repeat 15, without insert 2. (**C**) Two-axis views of the solvent accessible surface of insert 2. Note that the presence of the insert is predicted to cause a looping out of the adjacent helix A to helix B joining segment, creating a potential ligand attachment site. The new structure created by insert 2, which draws from sequences outside the insert, is amphipathic, in that it has one face largely hydrophobic, the other charged.

accord with current spectrin naming conventions (*3*, *32*, *33*), we term these isoforms of spectrin αIIΣ1, αIIΣ2, αIIΣ3, and αIIΣ4. The precise structure we have assigned to each of these transcripts is presented in Table 3. It is also surprising in this study that we did not detect using a sensitive PCR assay differential utilization of insert 3 in any tissue. This insert has only been identified in a single clone derived from lung fibroblasts (*15*), and the question remains open whether it represents a bona fide transcript. However, inspection of the genomic structure of the αI spectrin gene suggests that it is, since an almost identical sequence constitutes exon 46 of αI spectrin. If insert 3 in αII spectrin is real, it must represents a rare form of spectrin and possibly a cell-type specific transcript. We anticipate (but at this point cannot prove) that insert 3, when it is utilized, will also appear independently of the alternative splices yielding inserts 1 and 2. If so, this would mean that there are a total of eight different transcripts of αII spectrin. The presence of such isoform diversity may contribute to the differential tissue reactivity often observed with anti-spectrin antibodies, such as in studies of cardiac and skeletal muscle (*34*, *35*) or of neurons (*36*). It is also important to note that while this report identifies four different spectrin transcripts that appear in a tissue specific distribution, the steady-state abundance of the proteins derived from each transcript has not been measured. Thus, while αII spectrin (all isoforms) is a reasonably abundant protein in probably all cell types, it remains unknown if the various αII spectrin mRNA transcripts identified here are translated with equivalent efficiency or if their protein products are all equally stable. The answer to these questions will require isoform-specific antibodies. While the preparation of such antibodies is underway and preliminary results indicate that both insert 1 positive and insert 1 negative protein is present in at least some cells (Cianci and Morrow, unpublished observations), additional studies will be required to conclusively address these questions and to determine whether the protein products of insert 2 and insert 3 also reflect the abundance of their mRNAs.

A central question that remains is the functional consequences of αII-spectrin isoform diversity. Presumably, the

inserted sequences are altering the function or stability of the overall molecule. The structural modeling reported here for insert 2 and our preliminary unpublished observations that there is a ligand that specifically binds insert 2 in brain extracts lends credence to this hypothesis. Indeed, the predicted structure for repeat 15 suggests that complimentary sequence variations may even have arisen within helix C that act in concert with the connecting region between helices A and B (where insert 2 occurs) to accommodate a novel binding site or its ligand. The structure predicted for insert 2 demonstrates an exposed peptide loop with opposed hydrophobic and charged faces (Figure 6), reminiscent of the structure of highly antigenic epitopes that bind antibodies or the interaction site between p53 and the ankyrin-like p53 BP (*37*, *38*).

In future studies, it will be of interest to identify the nature of the brain and any muscle specific ligand(s) that interact with insert 2, as well as inserts 1 and 3. While the roles of inserts 1 and 3 remain unknown, these splices arise in regions of the molecule likely to be functionally active (cf. Figures 2 and 3). Thus, insert 1 flanks the SH3 domain of αII spectrin; the primary sites of calpain and caspase 3 cleavage; and the site of calcium dependent calmodulin binding (*39*−*41*). In preliminary studies using antibodies specific for insert 1, we have noted that insert 1 may contribute to the control of the intracellular compartmentalization of αII spectrin, and anticipate that this insert will turn out to be important for understanding the dynamics of αII spectrin assembly in vivo. (Cianci and Morrow, unpublished observations). Finally, insert 3 is found in repeat 21, a region involved in nucleating heterodimer assembly during spectrin synthesis (*42*), and also a region flanking two calcium-binding sites similar to those found in calmodulin (*43*) and implicated in the interaction of spectrin with NMDA receptor channels (*44*).

In summary, the results presented here establish that at least four distinct isoforms of human αII spectrin are expressed in various tissues and identify a novel human brain and muscle specific isoform. If additional complexity arises from alternative transcription at insert 3, which is likely although not detected in these studies, then eight different isoforms of αII spectrin arise from a single gene on chromosome 9. The structural consequences of the novel insert reported here have also been modeled; these studies suggest that insert 2 forms a binding site in repeat 15 for an unknown ligand that is presumably present in brain and muscle tissue. When these results are combined with the known transcriptional diversity of the *β*-spectrin gene family (reviewed in *3*, *9*, and *45*), and the observation that any *β*-spectrin can apparently combine with any α-spectrin to create a functional heterodimer, the combinatorial (and presumably functional) complexity of the spectrin family is significant. A challenge in future studies will be to understand the full complement of ligands that interact with spectrin, and how their multiple functionality's contribute to spectrin's role as a membrane organizing center at the plasma membrane and on internal membranes (*3*, *4*, *12*, *13*).

ACKNOWLEDGMENT

Dr. Bernard Forget is acknowledged for kindly providing a sample of the original lung fibroblast clone used to characterize that αII spectrin.

REFERENCES

1. Devarajan, P., and Morrow, J. S. (1996) in *Membrane Protein-Cytoskeleton Complexes: Protein Interactions, Distributions and Functions* (Nelson, W. J., Ed.) pp 97−128, Academic Press, New York.
2. Bennett, V., and Gilligan, D. M. (1993) *Annu. Rev. Cell. Biol. 9*, 27−66.
3. Morrow, J. S., Rimm, D. L., Kennedy, S. P., Cianci, C. D., Sinard, J. H., and Weed, S. A. (1997) in *Handbook of Physiology* (Hoffman, J., and Jamieson, J., Eds.) pp 485−540, Oxford, London.
4. deMatteis, M. A., and Morrow, J. S. (1998) *Curr. Opin. Cell Biol. 10*, 542−549.
5. Huebner, K., Palumbo, A. P., Isobe, M., Kozak, C. A., Monaco, S., Rovera, G., Croce, C. M., and Curtis, P. J. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3790−3793.
6. Winkelmann, J. C., Leto, T. L., Watkins, P. C., Eddy, R., Shows, T. B., Linnenbach, A. J., Sahr, K. E., Kathuria, N., Marchesi, V. T., and Forget, B. G. (1988) *Blood 72*, 328−334.
7. Leto, T. L., Fortugno-Erikson, D., Barton, D., Yang-Fong, T. L., Francke, U., Harris, A. S., Morrow, J. S., Marchesi, V. T., and Benz, E. J., Jr. (1988) *Mol. Cell. Biol. 8*, 1−9.
8. Chang, J. G., Scarpa, A., Eddy, R. L., Byers, M. G., Harris, A. S., Morrow, J. S., Watkins, P., Shows, T. B., and Forget, B. G. (1993) *Genomics 17*, 287−293.
9. Stankewich, M. C., Tse, W. T., Peters, L. L., Ch'ng, Y., Devarajan, P., Lux, S. E., and Morrow, J. S. (1998) *Mol. Biol. Cell 9* (Suppl.), 264a.
10. Bloch, R. J., and Morrow, J. S. (1989) *J. Cell Biol. 108*, 481−493.
11. Pumplin, D. W. (1995) *J. Cell Sci. 108*, 3145−3154.
12. Beck, K. A., Buchanan, J. A., Malhotra, V., and Nelson, W. J. (1994) *J. Cell Biol. 127*, 707−723.
13. Devarajan, P., Stabach, P. R., Mann, A. S., Ardito, T., Kashgarian, M., and Morrow, J. S. (1996) *J. Cell Biol. 133*, 819−830.
14. Sahr, K. E., Laurila, P., Kotula, L., Scarpa, A. L., Coupal, E., Leto, T. L., Linnenbach, A. J., Winkelmann, J. C., Speicher, D. W., Marchesi, V. T., Curtis, P. J., and Forget, B. G. (1990) *J. Biol. Chem. 265*, 4434−4443.
15. Moon, R. T., and McMahon, A. P. (1990) *J. Biol. Chem. 265*, 4427−4433.
16. Winkelmann, J. C., Costa, F. F., Linzie, B. L., and Forget, B. G. (1990) *J. Biol. Chem. 265*, 20449−20454.
17. Winkelmann, J. C., Chang, J. G., Tse, W. T., Scarpa, A. L., Marchesi, V. T., and Forget, B. G. (1990) *J. Biol. Chem. 265*, 11827−11832.
18. Amin, K. M., Scarpa, A., Winkelmann, J. C., Curtis, P. J., and Forget, B. G. (1993) *Genomics 18*, 118−125.
19. Lombardo, C. R., Weed, S. A., Kennedy, S. P., Forget, B. G., and Morrow, J. S. (1994) *J. Biol. Chem. 269*, 29212−29219.
20. Godi, A., Santone, I., Pertile, P., Devarajan, P., Stabach, P. R., Morrow, J. S., Di Tullio, G., Polishuck, R., Petrucci, T. C., Luini, A., and De Matteis, M. A. (1998) *Proc. Natl. Acad. Scie. U.S.A. 95*, 8607−8612.
21. Hyvonen, M., Macias, M. J., Nilges, M., Oschkinat, H., Saraste, M., and Wilmanns, M. (1995) *EMBO J. 14*, 4676−85.
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
23. Stabach, P. R., Cianci, C. D., Glantz, S. B., Zhang, Z., and Morrow, J. S. (1997) *Biochemistry 36*, 57−65.
24. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
25. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res. 12*, 387−395.
26. Yan, Y., Winograd, E., Viel, A., Cronin, T., Harrison, S. C., and Branton, D. (1993) *Science 262*, 2027−2030.
27. Zhang, Z., Weed, S. A., and Morrow, J. S. (1999) *FASEB J. 13*, A1394.

28. Wasenius, V. M., Saraste, M., Salven, P., Eramaa, M., Holm, L., and Lehto, V. P. (1989) *J. Cell Biol. 108*, 79−93.
29. Dubreuil, R. R., Byers, T. J., Sillman, A. L., Bar-Zvi, D., Goldstein, L. S., and Branton, D. (1989) *J. Cell Biol. 109*, 2197−2205.
30. Kotula, L., Laury-Kleintop, L. D., Showe, L., Sahr, K., Linnenbach, A. J., Forget, B., and Curtis, P. J. (1991) *Genomics 9*, 131−140.
31. Gallagher, P. G., Petruzzi, M. J., Weed, S. A., Zhang, Z., Marchesi, S. L., Mohandas, N., Morrow, J. S., and Forget, B. G. (1997) *J. Clin. Invest. 99*, 267−277.
32. Winkelmann, J. C., and Forget, B. G. (1993) *Blood 81*, 3173−3185.
33. Malchiodi-Albedi, F., Ceccarini, M., Winkelmann, J. C., Morrow, J. S., and Petrucci, T. C. (1993) *J. of Cell Science 106*, 67−78.
34. Shile, P. (1989) in *Pathology, School of Medicine* pp 186, Yale, New Haven.
35. Vybiral, T., Winkelmann, J. C., Roberts, R., Joe, E. H., Casey, D. L., Williams, J. K., and Epstein, H. F. (1992) *Cell Motil. Cytoskel. 21*, 293−304.
36. Clark, M. B., Ma, Y., Bloom, M. L., Barker, J. E., Zagon, I. S., Zimmer, W. E., and Goodman, S. R. (1994) *Brain Res. 663*, 223−236.
37. Gorina, S., and Pavletich, N. P. (1996) *Science 274*, 1001−1005.
38. Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., and Pavletich, N. P. (1996) *Science 274*, 948−953.
39. Harris, A. S., Croall, D. E., and Morrow, J. S. (1989) *J. Biol. Chem. 264*, 17401−17408.
40. Harris, A. S., Croall, D., and Morrow, J. S. (1988) *J. Biol. Chem. 263*, 15754−15761.
41. Wang, K. K. W., Posmantur, R., Nath, R., McGinnis, K., Whitton, M., Talanian, R. V., Glantz, S. B., and Morrow, J. S. (1998) *J. Biol. Chem. 273*, 22490−22497.
42. Speicher, D. W., Weglarz, L., and DeSilva, T. M. (1992) *J. Biol. Chem. 267*, 14775−14782.
43. Trave, G., Pastore, A., Hyvonen, M., and Saraste, M. (1995) *Eur. J. Biochem. 227*, 35−42.
44. Wechsler, A., and Teichberg, V. I. (1998) *EMBO J. 17*, 3931−3939.
45. Goodman, S. R., Zimmer, W. E., Clark, M. B., Zagon, I. S., Barker, J. E., and Bloom, M. L. (1995) *Brain Res. Bull. 36*, 593−606.

BI991458K