Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 5671–5675, September 1985
Biochemistry

Remarkable homology among the internal repeats of erythroid and nonerythroid spectrin

(cytoskeleton/cDNA cloning/gene structure and evolution)

CONNIE S. BIRKENMEIER*, DAVID M. BODINE*†, ELIZABETH A. REPASKY‡, DAVID M. HELFMAN§, STEPHEN H. HUGHES§¶, AND JANE E. BARKER*

*The Jackson Laboratory, Bar Harbor, ME 04609; ‡Roswell Park Memorial Institute, Buffalo, NY 14163; and §Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724

Communicated by Elizabeth S. Russell, May 10, 1985

ABSTRACT **A cDNA clone for nonerythroid α-spectrin was identified by direct immunological screening of a chicken smooth muscle cDNA library. A library prepared in the expression plasmids pUC8 and pUC9 was screened with an antiserum specific for chicken α-spectrin. Blots of poly(A)$^+$ RNA from various tissues of chicken and mouse show that the cDNA hybridizes to an 8-kilobase mRNA. The cDNA hybridizes to a single-copy sequence on Southern blots of chicken genomic DNA. The complete nucleic acid sequence of the clone has a single 1419-base open reading frame. The derived amino acid sequence is organized into two partial and three complete 106-amino-acid repeats that show homology to the repeats described for human erythroid α- and β-spectrin. Immunological and biochemical data indicate that chicken nonerythroid and human erythroid α-spectrin are two of the more widely diverged members of the spectrin family of proteins. In this respect, the degree of homology found between them was unexpected. Our data suggest a common evolutionary origin for these two α-spectrins and allow some predictions concerning spectrin gene structure.**

The membrane structural protein spectrin was once thought to be unique to erythrocytes (ref. 1; for reviews, see refs. 2–5). Recently membrane structural proteins related to erythroid spectrin have been identified in nonerythroid tissues (6–9). Biochemical and functional comparisons of these immunologically crossreacting proteins have defined them as isoforms of spectrin (for reviews, see refs. 10 and 11). The spectrin isoforms show tissue-specific distribution and differential expression during development (12–18). Spectrin proteins are heterodimeric, with α and β subunits ranging in size from 220 to 265 kDa. The β subunits are heterogeneous, with three avian and two mammalian tissue-specific β subunits known thus far (11). It appears that a common α subunit occurs in all avian tissues, while in mammals the erythroid α subunit is distinct from the nonerythroid α subunit (11, 12). The avian α subunit is immunologically closely related to the mammalian nonerythroid α subunit but only weakly related to the mammalian erythroid α subunit. The existence of multiple forms of spectrin with tissue-specific patterns of expression poses interesting questions about the number and structure of the genes that encode the spectrin proteins.

We have obtained and sequenced a chicken α-spectrin cDNA clone from smooth muscle mRNA. A mRNA of ≈8 kilobases (kb) is found on hybridization blots of poly(A)$^+$ RNA from chicken tissues. The derived amino acid sequence of the clone was compared to the known amino acid sequence of human erythroid spectrin (19–21). Speicher and Marchesi (21) reported that both subunits of human erythroid spectrin are organized into multiple internal repeats 106 amino acids long. Even though the mammalian erythroid α- and the avian α-spectrin are two of the more diverged spectrin proteins, we find striking conservation of the human erythroid repeats in our chicken α-spectrin clone. Southern blot analysis of cellular DNA indicates that our clone is present as a single-copy sequence in the chicken genome. The implications of these data for the structure and evolution of the spectrin gene family are discussed.

MATERIALS AND METHODS

Isolation of cDNA Clones. Preparation and screening of the cDNA expression library is described in detail elsewhere (22). Smooth muscle poly(A)$^+$ RNA used for cDNA synthesis was prepared from day 11 embryonic chicken stomachs and gizzards. Antiserum used to screen the library was raised in rabbit against one-dimensional gel-purified chicken erythroid α-spectrin. Preparation and specificity are as described previously by Repasky *et al.* (9). The specificity can be represented as follows: avian erythroid α-spectrin = avian nonerythroid α-spectrin > mammalian nonerythroid α-spectrin > mammalian erythroid α-spectrin. The antiserum does not crossreact with β-spectrin.

DNA Sequencing and Data Analysis. Appropriate restriction enzyme fragments of the spectrin cDNA clone were subcloned into M13 mp18 and mp19 and sequenced by the dideoxy chain-termination method (23). Sequence data were analyzed by using the computer facility and software described by Boguski *et al.* (24) or an IBM-PC/AT microcomputer and the Pustell DNA/protein sequence analysis software available from International Biotechnologies (New Haven, CT).

Blot Hybridization. Adult mice of the C57BL/6J strain from the Jackson Laboratory and day 11 or day 16 embryonic chickens from SPAFAS were tissue sources for RNA. Reticulocyte RNA was obtained from (WB/Re × C57BL/6J)F_1, hereafter referred to as WBB6F_1 hybrid mice homozygous for the hemolytic anemia mutation normoblastosis (*nb*) (25). We showed previously that poly(A)$^+$ RNA from reticulocytes of *nb/nb* WBB6F_1 animals directs the *in vitro* synthesis of immunoprecipitable α-spectrin (25). Total RNA was isolated by the guanidine·HCl method (26). Poly(A)$^+$ RNA was isolated by oligo(dT)-cellulose chromatography (27), fractionated by formaldehyde/agarose gel electrophoresis (28), and transferred to nitrocellulose (29). Hybridization was as described by Jahn *et al.* (30). The

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

Abbreviations: kb, kilobase(s); bp, base pair(s).

†Present address: Clinical Hematology Branch, Building 10, Room 7C103, National Institutes of Health, Bethesda, MD 20205.
¶Present address: Frederick Cancer Research Facility, P.O. Box B, Frederick, MD 21701.

hybridization probe was ^{32}P-labeled by using the T4 DNA polymerase labeling system from Bethesda Research Laboratories. The stringency of the hybridization and washing conditions were determined according to Shaw *et al.* (31).

Southern Blots. Genomic DNA was isolated from day 16 embryonic chicken erythrocytes (26). Restriction enzyme-digested DNA was electrophoresed in 0.8% agarose and blotted onto nitrocellulose (32). The hybridization protocol and preparation of the hybridization probe were as described for blot hybridization.

RESULTS

Isolation and Characterization of α-Spectrin cDNA Clones. Two positive clones detected by immunological screening of the cDNA library were identical by restriction enzyme mapping. One clone, pUC8-13a, was used for all subsequent experiments. Selection of clones by immunological screening of an expression cDNA library gave a high degree of confidence in the identity of the clone. The tissue of origin for the cDNA library was chicken smooth muscle, and the antiserum used for screening was specific for chicken α-spectrin. For these reasons we believe we have isolated a chicken α-spectrin cDNA clone, which we hereafter refer to as chicken nonerythroid α-spectrin to indicate the type of tissue it originated from and its immunological relationship to the mammalian nonerythroid α subunits.

Blot hybridization analysis supports the conclusion that pUC8-13a is a chicken α-spectrin cDNA clone. An appropriately sized mRNA (8 kb) to code for the 240-kDa α-spectrin was found in each tissue examined (Fig. 1*A*). Similarly sized mRNAs were found by Moon *et al.* using a chicken erythroid α-spectrin cDNA clone (33) and by Cioe and Curtis using a mouse erythroid α-spectrin cDNA clone (34). The intensity of hybridization to RNAs from the different tissues reflects the immunological crossreactivity of our antiserum (see *Materials and Methods*). Low-stringency hybridization conditions were needed to detect a band in mouse reticulocyte poly(A)$^+$ RNA (Fig. 1*B*). When analyzing RNA in the 8-kb size range, small differences in size are difficult to determine. However, we observed on numerous blots a slight size difference between the chicken erythroid mRNA (Fig. 1, lane 4) and the nonerythroid mRNAs (Fig. 1, lanes 1–3). This result suggests the possibility that more than one α-spectrin mRNA occurs in chickens.

DNA Sequence and Derived Amino Acid Sequence. Fig. 2 shows a partial restriction map of the cDNA insert and a summary of fragments sequenced. The nucleic acid sequence and derived amino acid sequence are shown in Fig. 3. Despite the fact that this clone expresses immunologically detectable α-spectrin, we were unable to match the reading frame of the vector with the reading frame of the insert. Sequence obtained around the 5′ cloning site was ambiguous even though we used several alternative sequencing methods, including Maxam and Gilbert (35) and dideoxy sequencing directly from the double-stranded pUC8-13a plasmid. We believe

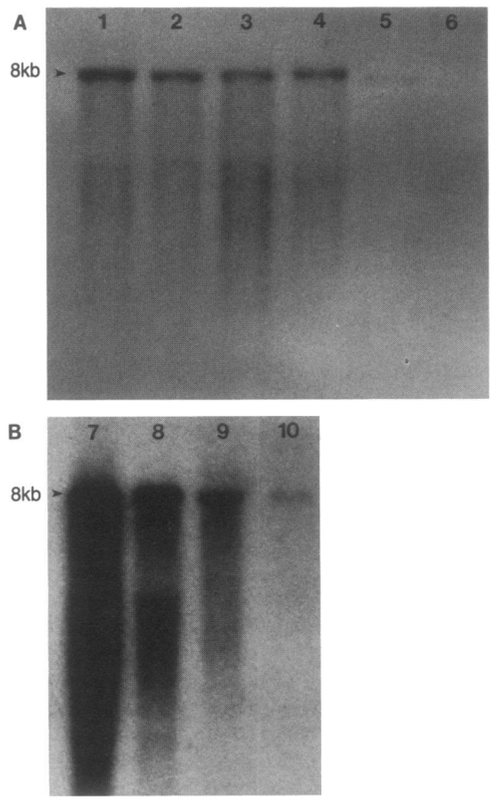

FIG. 1. Blot-hybridization analysis of poly(A)$^+$ RNA from various tissues of chicken and mouse. Lanes: 1–6, 5 μg of poly(A)$^+$ RNA isolated from day 11 embryonic chicken gizzard (lane 1), day 11 embryonic chicken stomach (lane 2), day 16 embryonic chicken brain (lane 3), day 16 embryonic chicken erythrocytes (lane 4), adult mouse brain (lane 5), adult mouse reticulocytes (WBB6F$_1$, *nb/nb*) (lane 6); 7–10, 10 μg of the same RNAs as used in lanes 3–6. The conditions of hybridization and washing were chosen so that a match of 88% in *A* and 65% in *B* was required for hybridization to occur. Fragments of *Hin*dIII-cut phage λ DNA were used as markers.

secondary structure in this region interfered with our ability to obtain accurate sequence. In our data analysis, we excluded the first 16 bases of the 5′ end of the insert from consideration.

Identification and Analysis of Homologous Repeat Structure. Using the single letter code (36), we aligned the derived amino acid sequence of pUC8-13a according to the convention Speicher and Marchesi used for human erythroid α- and β-spectrin (21). Two partial and three complete homologous repeats are apparent (Fig. 4). In pairwise comparisons, the amino acid homology among the chicken nonerythroid and human erythroid α-spectrin repeats is 20–36%. The homology is more impressive when one notes that certain patterns of amino acids are conserved. A consensus sequence of amino acids that predominate at certain positions in both the

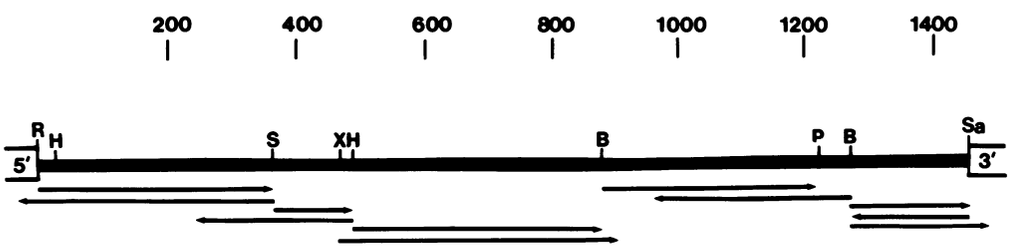

FIG. 2. Restriction map of α-spectrin clone pUC8-13a and summary of its sequence analysis. The scale is in nucleotides. The arrows represent the fragments sequenced and the direction of sequencing. R, *Eco*RI; H, *Hin*dIII; S, *Sst* I; X, *Xba* I; B, *Bgl* II; P, *Pst* I; Sa, *Sal* I.

```
        10         20         30         40         50         60         70         80         90        100        110        120
GAGGTGGAAGCTTTGTTGGCATCTGAAGACTATGGGAAGGACTTGGCATCAGTGAACAACCTTCTGAAGAAGCACCAATTACTGGAAGCTGATATATCTGCTCATGAGGATCGGCTGAAG
GluValGluAlaLeuLeuAlaSerGluAspTyrGlyLysAspLeuAlaSerValAsnAsnLeuLeuLysLysHisGlnLeuGluGluAlaAspIleSerAlaHisGluAspArgLeuLys

        130        140        150        160        170        180        190        200        210        220        230        240
GACCTGAACAGCCAGGCTGACAGTTTGATGACCAGCAGTGCTTTCGATACCTCCCAAGTAAAGGATAAACGTGAAACTATAAATGGGCGCTTCCAGAGAATCAAGAGCATGGCAGCTGCC
AspLeuAsnSerGlnAlaAspSerLeuMetThrSerSerAlaPheAspThrSerGlnValLysAspLysArgGluThrIleAsnGlyArgPheGlnArgIleLysSerMetAlaAlaAla

        250        260        270        280        290        300        310        320        330        340        350        360
CGCCGTGCGAAGCTCAACGAGTCGCACCGCTTGCATCAGTTCTTCCGTGACATGGATGATGGGGAGTCCTGGATCAAAGAGAAGAAACTGTTGGTGAGCTCAGAGGACTATGGCAGAGAC
ArgArgAlaLysLeuAsnGluSerHisArgLeuHisGlnPhePheArgAspMetAspAspGlyGluSerTrpIleLysGluLysLysLeuLeuValSerSerGluAspTyrGlyArgAsp

        370        380        390        400        410        420        430        440        450        460        470        480
CTGACTGGTGTGCAGAACCTGAGGAAGAAACATAAGCGCTTGGAAGCAGAATTAGCTGCCCATGAACCTGCTATCCAGGGTGTTCTAGACACTGGCAAGAAGCTTTCAGATGATAACACA
LeuThrGlyValGlnAsnLeuArgLysLysHisLysArgLeuGluAlaGluLeuAlaAlaHisGluProAlaIleGlnGlyValLeuAspThrGlyLysLysLeuSerAspAspAsnThr

        490        500        510        520        530        540        550        560        570        580        590        600
ATTGGGAAGGAGGAGATACAGCAGAGACTGGCTCGATTTGTGGACCACTGGAAAGAGTTAAAACAGCTGGCAGCTGCTCGGGGGCAGCGTCTGGAGGAGTCCCTGGAGTACCAACAGTTT
IleGlyLysGluGluIleGlnGlnArgLeuAlaArgPheValAspHisTrpLysGluLeuLysGlnLeuAlaAlaAlaArgGlyGlnArgLeuGluGluSerLeuGluTyrGlnGlnPhe

        610        620        630        640        650        660        670        680        690        700        710        720
GTAGCAAATGTTGAGGAAGAGGAGGCATGGATCAACGAGAAAATGACATTGGTAGCCAGTGAGGATTATGGAGACACACTTGCTGCTATCCAGGGCTTGCTGAAGAAGCATGAAGCATTC
ValAlaAsnValGluGluGluGluAlaTrpIleAsnGluLysMetThrLeuValAlaSerGluAspTyrGlyAspThrLeuAlaAlaIleGlnGlyLeuLeuLysLysHisGluAlaPhe

        730        740        750        760        770        780        790        800        810        820        830        840
GAGACTGACTTTACTGTCCACAAAGACAGAGTGAATGATGTTTGCGCTAATGGAGAGGATCTCATTAAAAAGAATAATCACCATGTGGAGAACATTACTGCTAAGATGAAGGGCCTCAAG
GluThrAspPheThrValHisLysAspArgValAsnAspValCysAlaAsnGlyGluAspLeuIleLysLysAsnAsnHisHisValGluAsnIleThrAlaLysMetLysGlyLeuLys

        850        860        870        880        890        900        910        920        930        940        950        960
GGCAAGGTATCAGATCTGGAGAAAGCCGCAGCTCAGAGGAAAGCCAAACTGGATGAGAACTCTGCCTTCCTCCAGTTCAACTGGAAAGCAGATGTGGTGGAGTCATGGATAGGTGAGAAG
GlyLysValSerAspLeuGluLysAlaAlaAlaGlnArgLysAlaLysLeuAspGluAsnSerAlaPheLeuGlnPheAsnTrpLysAlaAspValValGluSerTrpIleGlyGluLys

        970        980        990       1000       1010       1020       1030       1040       1050       1060       1070       1080
GAAAACAGTCTGAAGACAGATGATTATGGCCGTGACCTCTCCTCTGTGCAAACACTACTCACCAAACAGGAAACCTTTGATGCTGGACTTCAGGCTTTCCAGCAGGAGGGAATTGCAAAC
GluAsnSerLeuLysThrAspAspTyrGlyArgAspLeuSerSerValGlnThrLeuLeuThrLysGlnGluThrPheAspAlaGlyLeuGlnAlaPheGlnGlnGluGlyIleAlaAsn

       1090       1100       1110       1120       1130       1140       1150       1160       1170       1180       1190       1200
ATCACTGCTCTGAAAGACCAGCTACTGGCAGCCAAACATATCCAATCAAAGGCCATTGAGGTTCGTCACGCTTCCTTGATGAAACGCTGGAATCAGCTGCTTGCTAATTCTGCAGCCAGG
IleThrAlaLeuLysAspGlnLeuLeuAlaAlaLysHisIleGlnSerLysAlaIleGluValArgHisAlaSerLeuMetLysArgTrpAsnGlnLeuLeuAlaAsnSerAlaAlaArg

       1210       1220       1230       1240       1250       1260       1270       1280       1290       1300       1310       1320
AAAAAGAAACTCTTGGAGGCTCAGGAGCACTTCAGAAAGGTTGAAGATCTCTTCTTGACTTTTGCCAAGAAGGCCTCTGCCTTCAACAGTTGGTTTGAGAATGCTGAGGAGGACCTGACG
LysLysLysLeuLeuGluAlaGlnGluHisPheArgLysValGluAspLeuPheLeuThrPheAlaLysLysAlaSerAlaPheAsnSerTrpPheGluAsnAlaGluGluAspLeuThr

       1330       1340       1350       1360       1370       1380       1390       1400       1410       1420
GATCCCGTGCGCTGCAATTCCCTGGAAGAAATCAAAGCACTGCGAGAAGCCCACGATGCCTTCCGTTCCTCACTTAGCTCTGCCCAAGCTGACTTCAAC
AspProValArgCysAsnSerLeuGluGluIleLysAlaLeuArgGluAlaHisAspAlaPheArgSerSerLeuSerSerAlaGlnAlaAspPheAsn
```

FIG. 3. Nucleotide sequence and translation of chicken nonerythroid α-spectrin cDNA clone pUC8-13a. The presumptive *Eco*RI and *Sal* I linker sequences are excluded as well as the first 16 bases of the 5′ end of the insert (see *Results* for explanation).

chicken and human α spectrin repeats illustrates this point (line c in Fig. 4). The color coding emphasizes positions that are strongly conserved as to type of amino acid. In the data analysis shown in Fig. 5, the sequence homology was expressed in terms of correlation coefficients. According to the method of Kubota *et al.* (38), each amino acid was assigned a numerical value based upon a combination of physical and chemical parameters. This allowed conservative substitutions to be scored. The shorter 155-amino-acid chicken sequence was slid along the length of the human sequence, one residue moving at a time. At each position the correlation coefficient between the two sequences was plotted. A peak

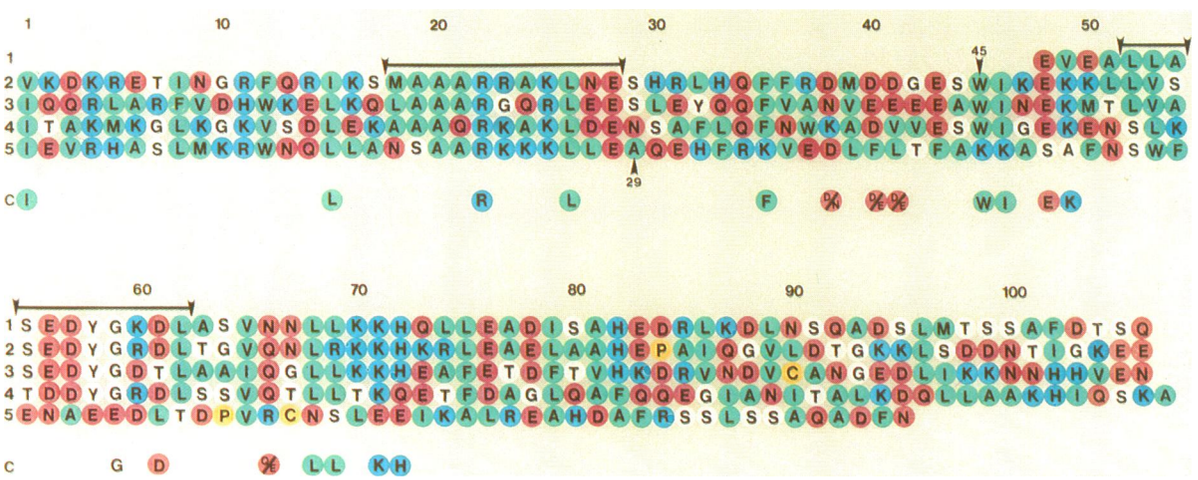

FIG. 4. Homologous repeat structure of chicken nonerythroid α-spectrin. The derived amino acid sequence is aligned as homologous repeating units 106 amino acids in length. The homologous repeats are lines 1–5 with amino acids 1–54 on the upper set and 55–106 on the lower set. Amino acids are color-coded based upon their hydropathy index and charge (24, 37). Green equals hydrophobic amino acids. Blue and red equal the positive- and negative-charged hydrophilic residues. The uncharged residues glutamine (Q) and asparagine (N) are also colored red because they have the same hydropathy indices (−3.5) as glutamic acid (E) and aspartic acid (D). Glycine (G), threonine (T), serine (S), and tyrosine (Y) have hydropathy indices near zero and are uncolored. Proline (P) and cysteine (C) are colored yellow to emphasize their unique structural properties. The bold line above the sequence demarks areas that are especially well conserved. The atypical region of repeat 5 becomes evident at residue 29 and is marked by an arrowhead. A consensus sequence labeled c is shown below repeat 5. It indicates amino acids that predominate in certain positions in both human and chicken repeats. The invariant tryptophane (W) residue at position 45 is indicated by an arrowhead.

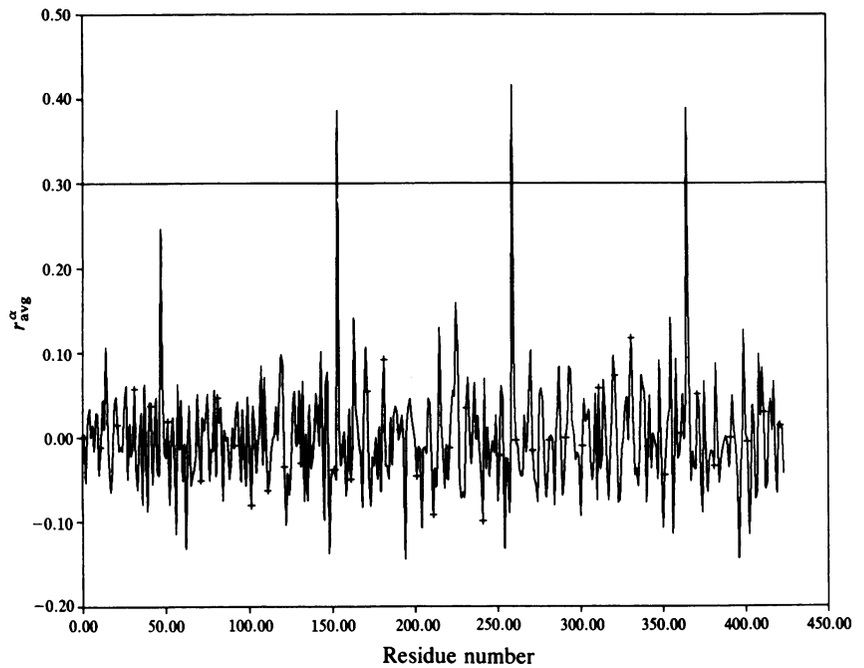

FIG. 5. Correlogram for the first 155 amino acids of the chicken nonerythroid α-spectrin cDNA clone compared to amino acids 18–595 from the αI domain of human erythroid α-spectrin (20). The average correlation coefficient, r_{avg}, is plotted along the y axis, and the residue number for the human sequence is plotted along the x axis. The horizontal line intersecting the y axis at 0.3 indicates a statistically significant match based upon results for proteins of known periodicity (38).

on the correlogram indicates a strong positive correlation and occurs when the chicken sequence is aligned with a highly homologous region of the human sequence. This occurs every 106 amino acids of the human sequence and confirms the relatedness of the chicken nonerythroid and human erythroid α-spectrin repeats.

Features of the Chicken Nonerythroid α-Spectrin Amino Acid Sequence. Analysis of the human erythroid α-spectrin sequence predicts that α-spectrin is composed of 20 homologous repeats, 106 amino acids in length, designated α-1 through α-20 (21). The NH_2-terminal sequence is the most complete, but partial sequence data exists for several repeats near the carboxyl terminus. Chicken repeat 4 (Fig. 4) is 51% homologous to the partial sequence (residues 24–53) of human α-18 repeat. As predicted by the method of cDNA preparation [oligo(dT) priming], pUC8-13a appears to be from the 3′ end of the mRNA. Chicken repeat 5 has a segment whose sequence is unrelated to the homologous repeat unit. Although it is difficult to determine precisely, homology seems to end at residue 29. Speicher and Marchesi have reported a similar loss of homology near the end of the α-9 repeat and throughout the α-10 repeat of human spectrin (21). There is not enough human α-spectrin sequence data available to determine if the human α-19 and α-20 repeats are atypical, but our data suggest that they will be.

Of special interest is the region from residue 55 to 62. This region is very strongly conserved in all but the atypical repeat of the chicken. It is also conserved in the α-3, α-7, and β-11 repeats of human spectrin. The region is very hydrophilic and is predicted to be exposed at the surface of the molecule, possibly serving as a binding or antigenic site.

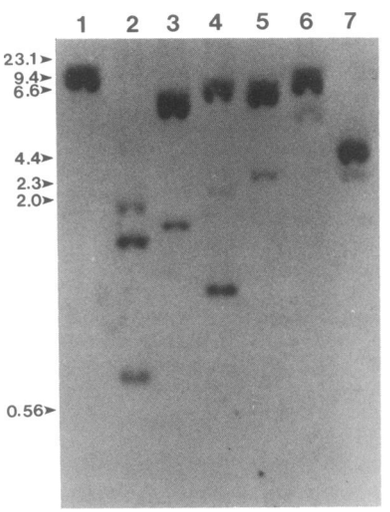

FIG. 6. Southern blot analysis of chicken genomic DNA. Each lane contains 10 μg of restriction enzyme-digested chicken genomic DNA. The blot was hybridized to ^{32}P-labeled cDNA clone pUC8-13a at 88% stringency. The size and position of *Hin*dIII-cut phage λ DNA used as markers are indicated to the left of the figure. Chicken genomic DNA was cut with *Eco*RI (lane 1), *Pst* I (lane 2), *Hin*dIII (lane 3), *Bgl* II (lane 4), *Sst* I (lane 5), *Bam*HI (lane 6), and *Bam*HI/*Sst* I (lane 7).

Southern Blot Analysis of Chicken Genomic DNA. Fig. 6 shows that clone pUC8-13a was detected as a single-copy sequence in the chicken genome when hybridized under conditions that require 88% matched bases to form a hybrid. A similar-length cDNA clone for chicken glyceraldehyde-3-phosphate dehydrogenase was used as a single-copy control for signal strength (data not shown). In preliminary experiments using lowered stringency conditions (70%), we detected additional fragments on Southern blots of chicken genomic DNA cut with these same restriction enzymes (data not shown). From the restriction map of the cDNA clone, a *Hin*dIII fragment of 452 base pairs (bp), a *Bgl* II fragment of 393 bp, and a *Bam*HI–*Sst* I fragment of 980 bp are predicted to occur in chicken genomic DNA. The genomic fragments detected at 88% stringency are all larger than predicted, suggesting that they contain introns.

DISCUSSION

Biochemical, immunological, and functional considerations suggest that, among the α-spectrin proteins, avian nonerythroid and mammalian erythroid are two of the more widely diverged isoforms. Yet comparison of the derived amino acid sequence of the chicken nonerythroid α-spectrin cDNA clone pUC8-13a to the amino acid sequence of human erythroid α-spectrin reveals striking similarity between these two molecules. The internal repeating structure first described by

Speicher *et al.* (19–21) for human erythroid spectrin is strongly conserved in chicken nonerythroid α-spectrin. When the human erythroid and chicken nonerythroid repeats are considered in pairwise comparisons, it is apparent that, while the amino acid homology (20–36%) has diverged considerably, there has been strong selection to conserve the structural potential of the repeats. The degree of similarity among the repeat units of these two sequences leaves little doubt that they arose from a common ancestral gene. The amino acid divergence among the repeats suggests that the repeating structure and possibly the genes for the various spectrin isoforms arose very early in evolution, probably before the divergence of the extant vertebrates.

The existence of a multigene family that includes genes for both the α and β subunits is highly likely. At the very least, the avian system suggests the possibility of four genes, one α-spectrin gene and three β-spectrin genes (β/β′,γ, and TW260) (11). In mammals the erythroid and nonerythroid α-spectrins are distinct (11), suggesting a minimum of two α-spectrin genes. Our cDNA clone occurs as a single-copy sequence in the chicken genome. However, this is not surprising when one considers that the nucleic acid homology among repeats is 45–50%. Even in the highly conserved SEDYG region of the repeat, because of third-base degeneracy, homology is only ≈75%. At this level of homology and under our stringent hybridization conditions, it is unlikely that the cDNA clone recognizes spectrin sequences other than itself in the genome. We predict that the next-most-homologous sequences to a given cDNA would be comparable regions of related genes rather than other regions of the same gene. The detection of additional fragments on Southern blots of chicken genomic DNA under lowered stringency conditions, as well as the mRNA size differences noted in Fig. 1, raise the possibility that multiple α-spectrin genes occur in the chicken. Furthermore, our data suggest that tissue-specific cDNA probes will be useful directly in hybridization experiments designed to define the number and chromosomal location of spectrin genes.

It is interesting to speculate about the gene structure and the nature of the ancestral gene that may have given rise to the spectrin gene family. The simplest hypothesis as suggested by Speicher *et al.* (19–21) is that the ancestral gene encoded one 106-amino-acid repeat and that the spectrin genes that exist today arose by a series of duplication and fusion events. If this were true, one might predict that gene structure, as reflected by intron–exon boundaries, would follow the homologous repeating structure seen in the protein sequence. Using a consensus sequence derived by Mount (39), we searched the nucleic acid sequence of clone pUC8-13a for potential intron–exon boundaries. Among the sites found, a set of four occur with a periodicity of 318 bp. These sites are positioned between amino acid residues 87 and 88 in each of the typical repeats represented by our clone, suggesting that spectrin gene structure may be related to the internal repeats of the spectrin proteins.

Another large structural protein, fibronectin (>200 kDa), shows intron–exon structure related to internal repeats of ≈150 nucleic acids. This entire gene has been isolated from chicken and is one of the largest genes known. It spans 48 kb and contains 48 exons (40). We predict that the spectrin genes will be of similar size.

Note Added in Proof. The cDNA clone 18-3a (41) is the same as our clone pUC8-13a. The sequence for clone 18-3a differs from ours. It lacks the *Hin*dIII site near the 5′ end of pUC8-13a. Discrepancies also exist in the region corresponding to nucleotides 952–969 of pUC8-13a and nucleotides 977–991 of 18-3a. The sequence of 18-3a requires gaps for optimal alignment of the protein repeats, while ours does not.

We thank Mark Boguski and Dr. Jeffery Gordon for their interest and help with the sequencing and computer analysis. We also thank Dr. Seldon Bernstein for his continued support and Drs. Edward Birkenmeier and Nancy Jenkins for their interest and help throughout this study. This work was supported in part by grants HL29305, AM27726, CA29569-02, and T32-CA09217-04 from the National Institutes of Health and a Muscular Dystrophy Postdoctoral Fellowship to D.M.H. The Jackson Laboratory is fully accredited by the American Association for Accreditation of Laboratory Animal Care.

1. Marchesi, V. T. & Steers, E. (1968) *Science* **159,** 203–204.
2. Branton, D., Cohen, C. M. & Tyler, J. (1981) *Cell* **24,** 24–32.
3. Goodman, S. R. & Shiffer, K. (1983) *Am. J. Physiol.* **244,** C121–C141.
4. Cohen, C. M. (1983) *Semin. Hematol.* **20,** 141–158.
5. Marchesi, V. T. (1983) *Blood* **61,** 1–11.
6. Goodman, S. R., Zagon, I. S. & Kulikowski, R. R. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 7570–7574.
7. Levine, J. & Willard, M. (1981) *J. Cell Biol.* **90,** 631–643.
8. Burridge, K., Kelly, T. & Mangeat, P. (1982) *J. Cell Biol.* **95,** 478–486.
9. Repasky, E., Granger, B. & Lazarides, E. (1982) *Cell* **29,** 821–833.
10. Bennett, V. (1984) in *Cell Membranes: Methods and Reviews*, eds. Elson, E., Frazier, W. & Glaser, L. (Plenum, New York), Vol. 2, pp. 149–195.
11. Nelson, W. J. & Lazarides, E. (1984) in *Cell Membranes: Methods and Reviews*, eds. Elson, E., Frazier, W. & Glaser, L. (Plenum, New York), Vol. 2, pp. 219–246.
12. Bennett, V., Davis, J. & Fowler, W. (1982) *Nature (London)* **299,** 126–131.
13. Glenney, J., Glenney, P. & Weber, K. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 4002–4005.
14. Nelson, W. J. & Lazarides, E. (1983) *Nature (London)* **304,** 364–368.
15. Glenney, J. R. & Glenney, P. (1983) *Cell* **34,** 503–512.
16. Lazarides, E. & Nelson, W. J. (1983) *Science* **222,** 931–933.
17. Lazarides, E., Nelson, W. J. & Kasamatsu, T. (1984) *Cell* **36,** 269–278.
18. Nelson, W. J. & Lazarides, E. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 363–367.
19. Speicher, D. W., Davis, G., Yurchenco, P. D. & Marchesi, V. T. (1983) *J. Biol. Chem.* **258,** 14931–14937.
20. Speicher, D. W., Davis, G. & Marchesi, V. T. (1983) *J. Biol. Chem.* **258,** 14938–14947.
21. Speicher, D. W. & Marchesi, V. T. (1984) *Nature (London)* **311,** 177–180.
22. Helfman, D. M., Feramisco, J. R., Ricci, W. M. & Hughes, S. H. (1984) *J. Biol. Chem.* **259,** 14136–14143.
23. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
24. Boguski, M. S., Elshourbagy, N., Taylor, J. M. & Gordon, J. I. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 5021–5025.
25. Bodine, D. M., Birkenmeier, C. S. & Barker, J. E. (1984) *Cell* **37,** 721–729.
26. Kozak, L. P. & Birkenmeier, E. H. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3020–3024.
27. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
28. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1983) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
29. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
30. Jahn, C. L., Hutchison, C. A., III, Phillips, S. J., Weaver, S., Haigwood, N. L., Voliva, C. F. & Edgell, M. H. (1980) *Cell* **21,** 159–168.
31. Shaw, G. M., Gonda, M. A., Flickinger, G. H., Hahn, B. H., Gallo, R. C. & Wong-Staal, F. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 4544–4548.
32. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
33. Moon, R. T., Ngai, J., Wold, B. J. & Lazarides, E. (1985) *J. Cell Biol.* **100,** 152–160.
34. Cioe, L. & Curtis, P. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1367–1371.
35. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
36. Hunt, L. T., Barker, W. C. & Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (National Biomedical Research Foundation, Washington, DC), Vol. 5, Suppl. 3, pp. 25–27.
37. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
38. Kubota, Y., Takahashi, S., Nishikawa, K. & Ooi, T. (1981) *J. Theor. Biol.* **91,** 347–361.
39. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–472.
40. Yamada, K. M. & Akiyama, S. K. (1984) in *Cell Membranes: Methods and Reviews*, eds. Elson, E., Frazier, W. & Glaser, L. (Plenum, New York), Vol. 2, pp. 77–148.
41. Wasenius, V.-T., Saraste, M., Knowles, J., Vertanen, I. & Lehto, V.-P. (1985) *Embo J.* **4,** 1425–1430.