Molecular Characteristics of the Novel Intermediate Filament Protein Paranemin

SEQUENCE REVEALS EAP-300 AND IFAPa-400 ARE HIGHLY HOMOLOGOUS TO PARANEMIN*

(Received for publication, July 28, 1997, and in revised form, September 16, 1997)

Philip M. Hemken, Robert M. Bellin, Suzanne W. Sernett, Bruno Becker, Ted W. Huiatt, and Richard M. Robson§

From the Muscle Biology Group, Departments of Biochemistry and Biophysics and of Animal Science, Iowa State University, Ames, Iowa 50011-3260

Paranemin was initially found to copurify with the intermediate filament (IF) proteins vimentin and desmin from embryonic chick skeletal muscle and was described as an IF-associated protein (IFAP). We have purified paranemin from embryonic chick skeletal muscle, prepared antibodies, and demonstrated that they label at the Z-lines of both adult avian and porcine cardiac and skeletal muscle myofibrils. We determined the cDNA sequence of paranemin by immunoscreening a agt22A cDNA library from embryonic chick skeletal muscle. Northern blot analysis revealed a single transcript of 5.5 kilobases, which is much smaller than predicted from the size of paranemin (280 kDa) by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The derived amino acid sequence of paranemin (1,606 residues; 178,161 kDa) contains the conserved IF rod domain (308 amino acids), which has highest homology to the rod domains of nestin and tanabin. Thus, paranemin is an IF protein rather than an IFAP. Sequence analysis also revealed that the partial cDNA sequences of two proteins, namely EAP-300 and IFAPa-400, are almost identical to regions of the cDNA sequence of paranemin. The complete paranemin cDNA was expressed in a cell line (SW13) with, and without, detectable cytoplasmic IFs. Antibody labeling of these cells suggests that paranemin does not form IFs by itself, but rather is incorporated into heteropolymeric IFs with vimentin.

Intermediate filaments (IFs),3 together with microfilaments and microtubules, comprise the three major classes of cytoskeletal filaments in cells of nearly all differentiated, multicellular eukaryotes (1–5). Much is now known about the cellular distribution, structure, and assembly of IFs (for recent reviews, see Refs. 3–6). Most IF proteins can be grouped into five major types or classes based upon their sequence and structure (1–3, 5), with a small number of IF proteins possibly comprising additional classes (3). The IFs, in general, are often considered to play an important role in the mechanical integration of cellular space (7–9), and their more specific cellular functions recently are becoming evident as well (3–5, 10).

Intermediate filaments in mature striated muscle cells are composed primarily of desmin (7, 11, 12). They are present in a collar-like arrangement around the myofibrillar Z-lines where they appear to connect adjacent myofibrils together, and possibly help link the peripheral layer of myofibrils to costameric sites along the muscle cell membrane (4, 13–16).

Paranemin, ~280 kDa by SDS-PAGE, was first identified in embryonic chick skeletal muscle (17). Paranemin has been considered an IF-associated protein (IFAP) (1, 18) because it copurified with the type III IF proteins vimentin and desmin from embryonic muscle (17), and colocalized with the major IF proteins vimentin and desmin at the periphery of avian myofibrillar Z-lines (17, 19). Its immunolocalization indicated a developmentally regulated expression in chick myogenic cells, and a more restricted expression in adult chicken muscle cells (17, 19).

As shown herein, paranemin contains the ~310-amino acid rod domain characteristic of IF proteins. Therefore, paranemin is a member of the IF protein superfamily rather than an IFAP. By sequence comparisons with other IF proteins, we have found that paranemin shares some homologies to other IF proteins, but significant differences as well. We also show in this report that regions of paranemin’s sequence are almost identical to the partial cDNA sequences reported for two other proteins, EAP-300 (20) and IFAPa-400 (21). The functional ability of paranemin to assemble into IFs was tested by cell transfection of the complete paranemin cDNA. Expression of paranemin cDNA in SW13 cell clones that either do, or do not, express vimentin, suggests that paranemin by itself is unable to form homopolymeric IFs, but that it can coassemble with vimentin into heteropolymeric IFs.

EXPERIMENTAL PROCEDURES

Purification of Paranemin from 14-day Embryonic Chick Skeletal Muscle—Approximately 100 g of thigh and breast muscles were dissected from 144 14-day-old chick embryos and homogenized in 235 ml of 130 mM KCl, 5 mM EDTA, 20 mM Tris-HCl, pH 7.5, as described (17). The resulting homogenate was centrifuged for 15 min at 20,700 × g. The supernatant was filtered through glass wool and centrifuged for 90 min at 125,000 × g. The resulting supernatant, referred to as crude paranemin, was separated into four equal volumes (~50 ml, ~10 mg/
Characterization of the Novel IF Protein Paranemin

**FIG. 1.** Western blot analysis of major fractions obtained during purification of paranemin. Monoclonal antibody 4D3 was used to identify paranemin by Western blot analysis throughout purification. Molecular mass markers (adult chicken cardiac myofibrillar proteins in kDa): myosin heavy chain (205), α-actinin (42), and tropomyosin (34), are indicated on the left. Paranemin (P) is indicated by the arrow on the right. A, SDS-PAGE analysis of overall paranemin purification, stained with Coomassie Blue; B, Western blot of a duplicate gel. Lane 1, total embryonic muscle homogenate; lane 2, crude paranemin; lane 3, gel filtration partially-purified paranemin pool; lane 4, hydroxyapatite-partially-purified paranemin pool; lane 5, DEAE-cellulose-purified paranemin pool (purified paranemin).

**FIG. 2.** Specificity of anti-paranemin monoclonal and polyclonal antibodies. A: a–d, Western blots of paranemin purified from skeletal muscle of 14-day embryonic chicks and purified synemin from avian gizzard smooth muscle were tested with monoclonal antibodies 4D3 (a), 4C7 (b), 3B12 (c), and polyclonal (d) anti-paranemin. e, whole tissue homogenate from skeletal muscle of 14-day embryonic chicks was tested with monoclonal antibody 4D3, which was used to screen the expression library. B, Western blots of two strips cut from the same lane of a sample of calpain-digested purified paranemin were tested with monoclonal antibodies 4D3 (a) and 4C7 (b). Large arrows indicate paranemin, with an apparent molecular mass of 280 kDa, and the small arrow indicates the position of synemin (230 kDa). Migration of molecular mass markers (number SDS-6H, Sigma) are at the left.

*phenylmethylsulfonyl fluoride, 20 mM Tris-HCl, pH 7.5, and loaded onto a 2.6 cm × 20-cm DEAE-cellulose (Whatman, DE-52) column previously equilibrated with the dialysis buffer. The column was eluted in the dialysis buffer with a linear 0–2 M NaCl gradient at a flow rate of 24 ml/h. Paranemin-containing fractions eluted from the column at ~1 M NaCl, and were pooled and dialyzed against column equilibration buffer. The sample of purified paranemin could be stored at ~70 °C for at least 4 weeks without detectable degradation.

Preparation and Characterization of Monoclonal Antibodies—Monoclonal antibodies were prepared following procedures described by Hemken et al. (22). Fusions with SP2/0-Ag14 (ATCC CRL 1581) myeloma cells were performed essentially according to published protocols (23). Enzyme-linked immunosorbent assays were performed to screen the monoclonals as described (22).

Preparation of Polyclonal Antibodies—Paranemin bands from SDS-PAGE gels, each containing ~1 mg of protein, were minced and homogenized in 1-ml aliquots of phosphate-buffered saline and emulsified with equal volumes of TiterMax (Vaxcell, Inc.) adjuvant. New Zealand White specific pathogen-free rabbits were injected with purified paranemin at several sites subcutaneously over the back and at one site intramuscularly in the thigh. The rabbits were boosted twice after the initial injection at 4–6-week intervals. Titters were determined by Western blotting.

Amino Acid Analysis and Sequencing—Purified paranemin was subjected to SDS-PAGE and electropherotically transferred to polyvinylidene difluoride membranes (Micron Separations) (24). For amino acid analysis, the 280-kDa paranemin band was excised from the washed blots and hydrolyzed in vacuo in 6 N HCl at 150 °C for 1, 2, and 3 h (25). Amino acid composition was determined with an Applied Biosystems Amino Acid Analyzer at the Iowa State University Protein Facility. For amino acid sequencing, deblocking using trifluoroacetic acid was necessary (26). The sample was analyzed with a 477A Protein Sequencer/120A Analyzer (Applied Biosystems Inc.).

*SDS-PAGE and Western Blotting—SDS-PAGE and Western blotting were performed essentially as described previously (22). Purified paranemin and synemin (27, 28), and paranemin that had been partially digested with purified m-calpain (0.9 units/mg for 10 min) in 2 mM urea, 5 mM CaCl₂, 10 mM β-mercaptoethanol, and 10 mM Tris-HCl, pH 7.4, were run on SDS-polyacrylamide gels and transferred to nitrocellulose (29, 30). The blots were probed with monoclonal 4D3, anti-paranemin.
Characterization of the Novel IF Protein Paranemin

Digoxigenin DNA Labeling and Detection Kit (Boehringer Mannheim). A total of 200 positive plaques were chosen, 100 from each probe. PCR was used to screen λ phage mixes of all 200 positive plaques for sequence overlapping clone 9 by using a gene-specific primer spanning paranemin cDNA positions 2130–2147 and the Agt11 Upstream Amplimer Primer (CLONTECH). Twenty clones that contained a 2.2-kb product by PCR screening were purified to homogeneity. Clones 3, 24, 89, and 169 were subcloned into pBluecript II SK(+) vector for sequencing.

Northern Blot Analysis—Northern blots were performed essentially as described (34, 35) on poly(A)+ RNA samples from skeletal muscle cells of 14-day embryonic chick with synthetic RNA markers (Life Technologies, Inc.) for size determination. Hybridization was performed overnight at 42 °C using 32P random prime-labeled cDNA probes directed to a specific activity between 1.7 × 10^6 and 1.7 × 10^7 Bq/ml. Unbound probe was removed by washing the filter at a final stringency of 0.2 × SSC and 0.1% (w/v) SDS at 50 °C for 30 min. Filters were exposed to Hyperfilm-MP (Amersham) for 3 days at −70 °C. Synthetic RNA markers were visualized by staining in 0.04% (w/v) methylene blue in 0.5 M sodium acetate, pH 5.2.

Expression of cDNA-encoded Proteins and Western Blotting—Purified phages of clone 9 were tested with monoclonal antibodies 4C7 and 3B12 culture supernatant) or polyclonal antibodies to paranemin (32). Ten positive plaques were chosen, 100 from each probe. PCR screening were purified to homogeneity. Clones 3, 24, 89, and 169 were subcloned into pBluecript II SK(+) vector for sequencing.

FIG. 5. Western blot analysis of the β-galactosidase-paranemin fusion protein. Control Y1089r− bacteria (lane 1, supernatant; lane 2, pellet) and Agt22A-clone 9 lysogenic Y1089r− bacteria (lanes 3, 4, and 5, supernatants) were subjected to SDS-PAGE (5% (w/v) stacking (S) gel and 8% (w/v) separating gel). Coomassie Blue (A) staining pattern and Western blot (B), using monoclonal antibody 4D3, of a duplicate gel transferred to nitrocellulose. Protein loads were approximately 20 μg in lane 1, 40 μg in lane 2, 5 μg in lane 3, 10 μg in lane 4, and 20 μg in lane 5. Arrows indicate the β-galactosidase-paranemin fusion protein. Lane M, molecular mass markers (Sigma) in kDa.

RNA Isolation and Construction of cDNA Library—Total RNA was isolated from skeletal muscle of 14-day embryonic chick by extraction with guanidinium thiocyanate at 4 °C (31). Poly(A)+ RNA was purified by chromatography on an oligo(dT)-cellulose column (Collaborative Biomedical Products) (32). Ten μg of poly(A)+ RNA were used to construct an oligo(dT)-primed, directional cDNA expression library in λgt22A using the Lambda SuperScript system (Life Technologies, Inc.).

Screening and Isolation of cDNA Clones—By immunoscreening of the λgt22A cDNA library using monoclonal antibody 4D3 to paranemin and the ProtoBlot detection system (Promega Corp.), 132 positive plaques were selected. Of these, 20 were purified to homogeneity. For sequencing, the two clones having the largest inserts, 9 (3.2 kb) and 40 (2.7 kb), were subcloned into pBluecript II SK(+) vector (Stratagene) (33). The cDNA library was rescreened by hybridization using a 205-bp BstXI restriction enzyme-generated cDNA probe from the 5′ end of clone 9 and a 2.2-kb cDNA probe generated by 5′-rapid amplification of cDNA ends (5′-RACE). For the 2.2-kb 5′-RACE probe, reverse-transcription of a poly(A)+ RNA sample was performed with a gene-specific primer spanning paranemin cDNA positions 2148–2165. PCR amplification of the target cDNA was performed using a paranemin-specific primer spanning cDNA positions 2130–2147 and the Anchor Primer (Life Technologies, Inc.). The λgt22A library was screened with both probes using a 2.2-kb cDNA probe generated by 5′-rapid amplification of cDNA ends (5′-RACE). For the 2.2-kb 5′-RACE probe, reverse-transcription of a poly(A)+ RNA sample was performed with a gene-specific primer spanning paranemin cDNA positions 2148–2165. PCR amplification of the target cDNA was performed using a paranemin-specific primer spanning cDNA positions 2130–2147 and the Anchor Primer (Life Technologies, Inc.).
FIG. 6. Nucleotide sequence of avian muscle paranemin cDNA and its deduced amino acid sequence. The positive numbers of nucleotides start at the first residue of the coding sequence. Amino acids are designated by the one-letter code. Bold regions represent sequence location of pseudo-heptad repeat sequence (also see Fig. 8; the repeat sequence extends slightly at both ends beyond the denoted amino acids, but only those in bold were considered in the discussion of this region of the sequence). Underlined regions represent the matching avian muscle paranemin sequence obtained by automated Edman degradation of peptides (parentheses indicate equivocal amino acid assignments). Square brackets denote the extent of the sequence reported (20) for EAP-300 (GenBank accession number X80877). Curly brackets denote the extent of the
Characterization of the Novel IF Protein Paranemin

Paranemin—The chromatographic purification of a representative (five total) paranemin preparation is shown in Fig. 1. A high molecular mass protein (280 kDa) was detected in the whole embryonic muscle homogenate (lane 1) and in the crude paranemin (lane 2), together with the IF proteins vimentin (54 kDa) and desmin (53 kDa), and many other proteins (17). Fractons eluting just after the void volume from the gel filtration column contained a complex of primarily paranemin, vimentin, and desmin (lane 3; vimentin and desmin are in same major band). Examination of negatively stained samples of these early fractions by electron microscopy revealed the presence of many long, ~8–12 nm diameter filaments with irregular surface contour (data not shown), in agreement with results of Breetler and Lazarides (17). Only the first few fractions of the first peak collected immediately following the void volume were used for further purification. Hydroxyapatite chromatography was very effective in removing from the gel filtration, partially purified paranemin, small amounts of several proteins with molecular masses near that of paranemin, and near those of vimentin and desmin (lane 4). DEAE-cellulos column chromatography removed both the remaining vimentin and desmin and a trace of actin before elution of the purified paranemin (lane 5). The pooled fractions of purified paranemin are shown in Fig. 1 (lane 5). Paranemin made by this procedure was also examined by SDS-PAGE with higher percent acrylamide gels, and showed no evidence of lower molecular mass contaminants. Although Breetler and Lazarides (17) indicated the paranemin in their paranemin-enriched fractions migrated as two closely spaced polypeptides of very similar molecular mass by SDS-PAGE, we usually observed only one paranemin band, and sometimes one major band with a small molecular mass by SDS-PAGE; we usually observed only one paranemin band, and sometimes one major band with a small molecular mass by SDS-PAGE.

Proteins were visualized by double label immunofluorescence with anti-paranemin polyclonal antibodies and anti-vimentin monoclonal antibody AMF-17b (developed by Dr. A. B. Fulton, obtained from the Developmental Studies Hybridoma Bank maintained by the Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, and the Department of Biological Sciences, University of Iowa, under contract NO1-HD-2-3144 from the National Institute of Child Health and Human Development). Transfected cells were grown on coverslips for 48 h and then fixed with cold methanol for 6 min at ~20 °C. Coverslips were blocked for 30 min with 5% goat serum before a 1-h (37 °C) incubation with a mixture of the primary antibodies, with dilutions in phosphate-buffered saline of anti-paranemin, 1:100, and anti-vimentin, 1:5. The coverslips were thoroughly washed in phosphate-buffered saline and incubated for 2 h (37 °C) with a mixture of fluorescein isothiocyanate-labeled goat anti-rabbit (1:200) and tetramethylrhodamine isothiocyanate-labeled goat anti-mouse (1:200). After incubation, the coverslips were again thoroughly washed in phosphate-buffered saline, mounted on glass slides with FITC-Guard (Testog), and observed on a Zeiss Photomicroscope III.

RESULTS

Purification of Paranemin—The chromatographic purification of a representative (five total) paranemin preparation is shown in Fig. 1. A high molecular mass protein (280 kDa) was detected in the whole embryonic muscle homogenate (lane 1) and in the crude paranemin (lane 2), together with the IF proteins vimentin (54 kDa) and desmin (53 kDa), and many other proteins (17). Fractons eluting just after the void volume from the gel filtration column contained a complex of primarily paranemin, vimentin, and desmin (lane 3; vimentin and desmin are in same major band). Examination of negatively stained samples of these early fractions by electron microscopy revealed the presence of many long, ~8–12 nm diameter filaments with irregular surface contour (data not shown), in agreement with results of Breetler and Lazarides (17). Only the first few fractions of the first peak collected immediately following the void volume were used for further purification. Hydroxyapatite chromatography was very effective in removing from the gel filtration, partially purified paranemin, small amounts of several proteins with molecular masses near that of paranemin, and near those of vimentin and desmin (lane 4). DEAE-cellulos column chromatography removed both the remaining vimentin and desmin and a trace of actin before elution of the purified paranemin (lane 5). The pooled fractions of purified paranemin are shown in Fig. 1 (lane 5). Paranemin made by this procedure was also examined by SDS-PAGE with higher percent acrylamide gels, and showed no evidence of lower molecular mass contaminants. Although Breetler and Lazarides (17) indicated the paranemin in their paranemin-enriched fractions migrated as two closely spaced polypeptides of very similar molecular mass by SDS-PAGE, we usually observed only one paranemin band, and sometimes one major band with a small trace of a band migrating just below it, which also labeled with monoclonal antibody 4D3 by Western blotting. The yield of paranemin from 100 g of embryonic skeletal muscle ranged from 1.2 to 2.5 mg with an average of 1.7 mg. Paranemin made by this procedure was also examined by SDS-PAGE with higher percent acrylamide gels, and showed no evidence of lower molecular mass contaminants. Although Breetler and Lazarides (17) indicated the paranemin in their paranemin-enriched fractions migrated as two closely spaced polypeptides of very similar molecular mass by SDS-PAGE, we usually observed only one paranemin band, and sometimes one major band with a small trace of a band migrating just below it, which also labeled with monoclonal antibody 4D3 by Western blotting. The yield of paranemin from 100 g of embryonic skeletal muscle ranged from 1.2 to 2.5 mg with an average of 1.7 mg.

Paranemin Antibody Characterization and Immunolocalization—Two monoclonal antibodies, 4D3 and 4C7, and the rabbit polyclonal antibody reacted specifically with paranemin, which migrated at an apparent molecular mass of 280 kDa (Fig. 2); however, monoclonal antibody 3B12 recognized both paranemin and synemin (Fig. 2A). The monoclonal antibody 4D3 was tested against the 14-day embryonic chick skeletal whole mus-
Isolation of cDNA Clones—A cDNA expression library in Agt22A was constructed from poly(A)^+ RNA isolated from skeletal muscle of 14-day embryonic chicks using an oligo(dT)- primer. This library, which contained about 2.6 × 10^6 independent plaque forming units, was screened for the expression of paranemin epitopes by using monoclonal antibody 4D3. This antibody detected only a band corresponding to paranemin’s size on a Western blot of whole muscle extract from skeletal muscle of 14-day embryonic chicks (Fig. 2A). Out of 2.6 million plaques immunoscreened, 132 gave a positive signal. Plaques from the two plaques with the strongest signal were further purified by rescreening. The two largest cDNA inserts, of clones 9 (3.2 kb) and 40 (2.7 kb) (Fig. 2A), were then sequenced. For further screening, 5’-RACE was initially used to generate a 2.2-kb probe by using a primer at the 5’-end of clone 9. Hybridization screening, using this probe and a 205-bp BstXI probe (Fig. 4A) generated from the 5’-end of clone 9, yielded approximately 200 positive plaques. PCR analysis of the 200 phage mixes identified 20 clones, all of which extended 2.1 kb beyond the 5’-end of clone 9 and yielded the same size amplification (2.2 kb) as the 5’-RACE product. These 20 clones were purified to homogeneity and, of these, clones 3, 24, 89, and 169 were used for sequencing. Alignment of all sequenced clones is shown in Fig. 4A.

Northern Analysis—Northern blot analysis was used to determine the message size (Fig. 4B). As probes, we used two restriction enzyme cDNA fragments of clone 9, one (HaeIII fragment) is located in the region that we identified by BLAST searches to overlap with the cDNAs for EAP-300 (20) and IFAPa-400 (21), and the other one (BstXI fragment) is derived from the 5’-end that does not have any overlap to those partial cDNAs. The size of the mRNA for paranemin from embryonic chicken skeletal muscle was estimated to be 5.3 kb (Fig. 4B).

Characterization of Paranemin Clones—To confirm that the phages of clone 9 were indeed expressing cDNA encoding for paranemin, we tested blots of plaque plaques of clone 9 with a panel of monoclonal antibodies and rabbit polyclonal antibodies to paranemin. Monoclonal antibodies, 4D3 and 4C7, and rabbit polyclonal antibodies clearly bound to proteins contained in plaques of clone 9. To further prove plagues of clone 9 coded for paranemin, the β-galactosidase fusion protein was expressed in E. coli Y1089r− and lysates were analyzed on a Western blot (Fig. 5). The monoclonal antibody 4D3 specifically labeled a band of 220 kDa corresponding to the fusion protein (Fig. 5B, lanes 3–5), and no labeling was seen in lanes containing wild type Y1089r− bacterial proteins (control; Fig. 5B, lanes 1 and 2). Assuming an average molecular mass of 110 daltons for an amino acid residue in paranemin, we calculated from the size of the fusion protein that the cDNA insert of clone 9 contained an open reading frame equivalent to about 950 amino acids (~105 kDa). This is slightly less than the length predicted from the size of the cDNA insert (3.2 kb), indicating the presence of a short noncoding region on this clone, which was confirmed by sequencing results.

After obtaining cDNA clones that appeared to contain the 5’-end of paranemin by hybridization screening, PCR with the paranemin cDNA pBluescript II SK(+) clones 3 and 24 as the template was compared with 5’-RACE using mRNA as template with the same gene-specific primers to confirm that these clones contained the 5’-end of paranemin. In two comparisons for each clone, the 5’-RACE and PCR products were almost the same size, with the 5’-RACE product slightly smaller, demonstrating that both clones contained the 5’-end of paranemin. The slight difference in size was due to the amplification of 90 bp of the multiple cloning site of pBluescript II SK(+) vector for the PCR of clones 3 and 24, compared with only approximately 36 bp for the 5’-RACE of a modified anchor primer, which was synthesized without a 12-bp cloning site at its 5’-end.

To confirm that the entire coding region of the paranemin cDNA had been recovered, the size of the probe coded for by the complete transcript was determined by bacterial expression of a clone 24-pProExHTb construct. The resulting, expressed protein co-migrated with purified muscle paranemin at ~280 kDa, and was labeled by Western blotting with the paranemin polyclonal antibodies. These results (not shown) indicated that the recovered clones code for the entire protein sequence of paranemin.

Sequence Analysis—The complete nucleotide sequence and derived amino acid sequence of paranemin are shown in Fig. 6. A striking pseudo-heptad repeat region is located near the center of the molecule, from approximately paranemin nucleotide positions 2257 to 3033 (see bold sequence in Fig. 6, also see Fig. 8). The peptide sequence of SQEEHDLQVEHRYLRV obtained by sequencing of purified paranemin, and originally expected to be N-terminal amino acid sequence, was found at residues 963–979 with exact agreement at 13 of 17 positions (see underlined sequence in Fig. 6). At all four positions of mismatch the residue predicted by the cDNA sequence was the second most abundant amino acid detected in the sequence analysis. During the deblocking procedure (26), the paranemin molecule was cleaved in the pseudo-heptad repeat region and, as a result, internal sequence was obtained. The internal cleavage of the protein is a likely cause of the sequence mismatch at the four positions noted, because some cleavage at alternate sites would result in interfering products during protein sequence analysis. A comparison of the predicted and measured amino acid compositions (Table I) indicates close agreement, especially Glx, Ser, His, Arg, Thr, Val, Ile, Leu, Phe, and Lys. The Met, Cys, and Trp composition had not been experimentally determined. The calculated pI and molecular weight of paranemin, based on the cDNA sequence, are 4.17 and 178,161, respectively.

Alignment of the derived amino acid sequence of paranemin’s rod domain to those of the rod domains of seven other IF proteins chosen to represent other types of IF proteins is shown in Fig. 7. The overall sequence identity within the rod domain was determined to be in the range of 63.3% (tanabin) to 23.7% (keratin 14). The highest degree of conservation was found at both ends of the rod domain, as expected (3). At the extreme C terminus of the rod domain, the paranemin amino acid sequence completely matched the IF protein consensus (signature) sequence. Valine in the first position of the signature pattern is found only in paranemin, tanabin, and nestin, whereas isoleucine is found in all other types (Fig. 7).
The overall amino acid identities between paranemin and nestin, which has been referred to as a class VI IF protein (40), and between paranemin and tanabin, which has been proposed as a class VII IF protein (39), are 25.2 and 27.4%, respectively. The percent identity of the tail domain of paranemin to the tail domains of these two proteins is significantly lower than of the rod domains, with 48.5 (rod) and 21.8% (tail) between paranemin and nestin, and 63.3 (rod) and 17.7% (tail) between paranemin and tanabin. The N-terminal head domain of paranemin (15 residues) also was much shorter than those of most IF proteins (for reviews, see Refs. 1 and 3), but was similar in length to the short N-terminal head domains of nestin (7 residues (40)) and tanabin (12 residues (39)).

The comparison of the full-length cDNA sequence of paranemin (5.3 kb) with partial sequences for EAP-300 (1.4 kb) (20) and IFAPα-400 (1.7 kb) (21), found by BLAST searches of GenBank, revealed that paranemin and EAP-300 share 476/481 amino acids (99.0% identity) and 1436/1443 nucleotides (99.5% identity), whereas paranemin and IFAPα-400 share 428/431 amino acids (99.3% identity) and 1708/1711 nucleotides (99.8% identity) in their overlapping regions (see Figs. 4 and 6).

Secondary structural prediction models (45, 46) indicate paranemin is predominantly α-helical with some turns and β-sheets between the α-helices. The paranemin sequence is very hydrophilic, reflecting the acidic character of the sequence, which has a predicted isoelectric point of 4.17. The comparison of the full-length paranemin cDNA sequence with itself shows a major block of about 36 consecutive pseudoheptad repeats (Fig. 8), with the most common being LQEEHGD, LQVEHGD, LQVEHED, and TQEEHGD (see Fig. 6).

**DISCUSSION**

Breckler and Ladriere (17) reported that paranemin copurified with the IF proteins vimentin and desmin, but did not chromatographically separate paranemin from those two IF proteins or other contaminants. We prepared a high-speed min polyclonal antibodies in either SW13Cl.1Vim<sup>+</sup> or SW13Cl.2Vim<sup>-</sup> cells.

** FIG. 7. Optimal alignment of the rod domain of paranemin with the rod domains of other IF proteins.** Sequence identity to paranemin is indicated by dashes, and dots represent gaps inserted by the alignment program PILEUP with the GapWeight 5 and GapLength 5 0.1. Hydrophobic residues in the heptad repeat are indicated by asterisks. The start of helical domain 2A for paranemin was difficult to define. The regular pattern of heptad repeats is interrupted once in domain 2B by the presence of a “stutter.” The intermediate filament signature (double underline), [I/V]<sub>1</sub>[T/A/C/I]<sub>2</sub>[R/K/H]<sub>3</sub>[X]<sub>4</sub>[L/M]<sub>5</sub>[D/E], where <sup>a</sup>X can be any amino acid, was identified in paranemin with zero mismatches using the GCG program MOTIFS. The percent identity was calculated by individual alignments of each rod domain using the program GAP with the same GapWeight and GapLength as used with PILEUP. The 42-amino acid insertion in region 1B of the lamin A sequence was not included in the calculation of percent identity.

![Helical Domain 1A](helical_domain_1a.png)

![Helical Domain 1B](helical_domain_1b.png)

![Helical Domain 2A](helical_domain_2a.png)

![Helical Domain 2B](helical_domain_2b.png)

**TABLE 1**

| Amino acid | Predicted Mole % | Measured Mole % |
|------------|------------------|-----------------|
| Asp        | 6.102            | 7.721           |
| Asn        | 1.245            | 2.666           |
| Asx        | 7.347            | 10.1            |
| Glu        | 18.453           | 16.7            |
| Gln        | 7.410            | 2.366           |
| Glnx       | 25.847           | 25.1            |
| Ser        | 6.974            | 1.171           |
| Gly        | 9.407            | 8.904           |
| His        | 3.736            | 3.381           |
| Arg        | 4.981            | 3.612           |
| Thr        | 4.172            | 0.996           |
| Acidic     | 33.188           | 32.1            |

* ND, not determined.

![Characterization of the Novel IF Protein Paranemin](characterization_of_paranemin.png)
supernatant fraction (crude paranemin) from the same tissue, 14-day embryonic chick skeletal muscle, and a subsequent gel filtration partially-purified paranemin fraction, according to their procedures (17) to be certain we were purifying the same protein, paranemin. The subsequent chromatographic purification steps used herein were successful in producing highly purified protein for production of the antibodies used for many of the studies described, including confirmation that the cloned cDNA was that of paranemin.

We identified the paranemin cDNA clones by immunological screening of a phage expression library using monoclonal antibody 4D3. The identity of the paranemin sequence was further established by amino acid sequencing, Western blotting of a β-galactosidase fusion protein, and comparison of the predicted and measured amino acid compositions of purified paranemin.

An unexpected discovery of our study was that paranemin contains the rod domain characteristic of all cytoplasmic IF proteins and, therefore, it is an IF protein rather than an IFAP, its previous classification (1, 18). Of the IF protein superfamily, paranemin shares significant homology with human nestin (40) and frog tanabin (39) within the IF rod domain. Paranemin also shares in common with nestin and tanabin short N-terminal head domains and long C-terminal tail domains. However, neither nestin nor tanabin contain the long, unique stretch of pseudo-heptad repeats present in the C-terminal tail domain of paranemin. Because we have not cloned the genomic DNA for paranemin, it is premature to classify it into one of the pre-existing classes of IF proteins, or to establish a new class.

Sequence comparisons provide strong evidence that both EAP-300 (20) and IFAPa-400 (21) are highly homologous, if not identical, to paranemin (51). The near 100% identity between the partial sequences of EAP-300 and IFAPa-400 was reported by Kelly et al. (20). Both partial sequences are entirely included in the complete paranemin sequence. Because paranemin was named by Breckler and Lazarides in 1982 (17) previous to the naming of EAP-300 in 1992 by McCabe et al. (52) and IFAPa-400 by Vincent and Lahaie in 1988 (53), we suggest that, unless further studies yield evidence of discrete differences between the proteins, these latter two proteins be referred to as paranemin.

Simard et al. (21) suggested that the size of the single IFAPa-400 message, which they found in embryonic chick heart, skeletal muscle, and brain, was greater than 10 kb, but it was based on only two ribosomal RNA markers. We find the size of the single 14-day chick skeletal muscle transcript in Northern blots is 5.3 kb, when compared with six synthetic RNA markers. Furthermore, we used a probe (HaeIII) (Fig. 4B) that overlaps with the sequence used as the Northern probe for the IFAPa-400 (21). The 5.3-kb message size of the paranemin transcript is only large enough to code for an ~190-kDa protein, which is much less than the ~280 kDa estimated for paranemin by SDS-PAGE (studies herein, and Ref. 17), or the ~300 kDa estimated for EAP-300. Completion of the sequence revealed that the paranemin cDNA encodes a protein with a predicted molecular mass of ~178 kDa, in close agreement with
the transcript size revealed by the Northern blot.

The estimated sizes by SDS-PAGE of paranemin (280 kDa) and EAP-300 (300 kDa) are in good agreement, in comparison to the somewhat larger size estimated for IFAPα-400 (400 kDa). The aberrant migration of paranemin (also EAP-300 and IFAPα-400) in SDS gels in comparison to its size estimated from the sequence may be due to the large number of negatively charged residues (Table I), which may cause SDS to bind poorly to proteins (54). Other proteins, such as caldesmon (55) and calpastatin (56), also are rich in acidic amino acids and exhibit anomalous behavior by SDS-PAGE. The estimated molecular masses (in daltons), for instance, determined by SDS-PAGE for smooth muscle caldesmon (120,000 (57,58); 125,000 (59); 140,000 (60); and 150,000 (61)), are much higher than the estimated molecular mass based on the cDNA-derived sequence (87,000 (62)). And, the molecular mass of the large neurofilament subunit H determined by SDS-PAGE is ~220,000 daltons, but its predicted molecular mass from amino acid sequence is 112,000 daltons (63). That the full-length paranemin clone 24-pProExH1Tb construct expressed in bacteria, where very little, if any, post-translational modification occurs, resulted in a protein that co-migrated with paranemin purified from muscle demonstrated that paranemin’s aberrant migration at ~280 kDa is an inherent property of its primary sequence.

The long consecutive pseudo-heptad repeat region in paranemin’s tail domain is unique (51), not having been identified in any of the ~60 proteins in the IF superfamily. The sequence reveals that many of the seven individual residue positions within consecutive heptads are occupied by the same amino acid residue (i.e., 24/36 heptads contain Gln in position 1, 31/36 heptads contain Glu in position 2, 24/36 heptads contain Asp in position 3, 34/36 heptads contain Asp in position 4, 31/36 heptads contain Gly in position 5, 27/36 heptads contain Gly in position 6, and 34/36 heptads contain Asp in position 7). Secondary structure predictions indicate that the repeat region would be α-helical. It was proposed by Kelly et al. (20) that the portion of this heptad repeat region included in the partial sequence they reported for EAP-300 contains multiple leucine zipper-like motifs, which they suggest may play a role in protein-protein interactions.

Localized paranemin by Price and Lazarides (19) indicated that paranemin was expressed in all myogenic muscle tissues, but only remained in heart muscle and vascular muscle of elastic vessels of the adult chicken. We, however, observed some paranemin present at the myofibrillar Z-lines of both isolated myofibrils (Fig. 3) and frozen sections from adult chicken skeletal muscle. Differences in the results with adult skeletal muscle possibly can be explained by the fact that different antibodies and labeling techniques were used in the two studies. EAP-300 (52, 64) and IFAPα-400 (65) have been localized in embryonic chick skeletal and cardiac tissue and in adult cardiac tissue. In addition, embryonic chick neuronal tissues shown to express EAP-300 (52, 64, 66) and IFAPα-400 (67) include the brain, spinal cord, and peripheral nervous system. Because we have shown that EAP-300 and IFAPα-400 are very likely identical to paranemin, paranemin also may fulfill functions previously ascribed to those two proteins, although differences in post-translational modification in the non-muscle tissues are possible. The labeling of paranemin at the Z-lines of porcine cardiac and skeletal muscle myofibrils (Fig. 3) also indicates that at least an antigeneically-related homologue exists in mammalian species.

Because paranemin contains an IF rod domain, it is plausible that this region interacts with the rod domains of major IF proteins such as vimentin or desmin (3, 68) in the backbone or core of IFs, and that it links the resulting heteropolymeric IFs to other cytoskeletal components via its long C-terminal tail domain. Duval et al. (69) have shown blot overlay studies indicating that IFAPα-400 interacts with vimentin and desmin. The ability of paranemin to assemble with other IF proteins is supported by the results of the SW13 transfection studies (Fig. 9). These experiments are similar in design to studies conducted by other investigators examining assembly of other IF proteins, such as peripherin (70) or specific neurofilament protein subunits (71), into IFs. The SW13 cells are known as an IF mosaic cell line because labeling with antibodies against vimentin showed that not all cells contained a cytoplasmic vimentin network (72). The SW13 cell line was subsequently cloned into a number of clonal lines for the purpose of establishing a vimentin-free cell line (50). Whereas clone 1 cells express vimentin, and display a normal IF network, clone 2 cells do not express vimentin and do not contain IFs (50). Thus, use of these
clonal lines permits examination of transfection-based assembly of IF proteins, or mutants, in both the presence and absence of an endogenous IF network. We found in clone 1 cells that the expressed paranemin co-localizes with the vimentin IF network, indicating that it either is incorporating into and/or attaching to the IF network. In clone 2 cells, the expressed paranemin appears in a nonfilamentous, punctate distribution in the cytoplasm. A similar nonfilamentous, punctate labeling pattern was noted as evidence of lack of IF assembly by desmin deletion mutants in MCF-7 cells, and other vimentin-free cell line (73). The nonfilamentous labeling pattern observed with paranemin expressed in vimentin-free cells suggests that paranemin is unable to form an IF network without a major IF protein, such as vimentin, and thus must incorporate into heteropolymeric IFs. An analogous situation occurs with the two large neurofilament triplet proteins NF-H and NF-M, which only can assemble into IFs with the smaller triplet protein NF-L (74, 75). Our in vivo assembly studies with paranemin also are in concert with negative staining observations that paranemin by itself does not assemble into 10-nm diameter filaments when dialogized into IF-forming buffer.

In summary, the sequence revealed several unexpected and significant properties of paranemin, including: 1) paranemin is an IF protein rather than an IFAP; 2) paranemin is a huge (178 kDa) IF protein, being much larger, for instance, than even neurofilament-H (~112 kDa), which is the largest of the triplet protein subunits composing neurofilaments; 3) paranemin shares several molecular features in common with the IF proteins nestin and tanabini, but significant differences as well; 4) paranemin contains approximately 36 consecutive, unusual pseudo-heptad repeats near the middle of the molecule, which have not previously been described in any of the over 60 IF proteins; and 5) regions within the complete sequence of paranemin are nearly identical to the partial sequences reported for two other proteins, EAP-300 and IFAPa-400. Transfection studies using SW13 cells have shown that expressed paranemin only is present in IFs in the presence of a major IF protein. Thus, paranemin likely functions within cells as a component of heteropolymeric IFs. As an IF component, paranemin should have an important role in modulating IF function in developing and mature cardiac and skeletal muscle cells and in other cells such as developing neurons. Taken in toto the results herein provide important information to help discover how the IFs in cells of muscle, nerves, and possibly other tissues, are linked to other cytoskeletal structures and regulate aspects of the cell cytoskeleton.

Acknowledgments—We thank Dr. Wallace Ip, University of Cincinnati School of Medicine, for providing the SW13 cell lines, Dr. Elisabeth Huff-Lonergan, Auburn University, for advice regarding immunolocalization studies, Dr. Elisabeth Huff-Lonergan, Auburn University, for providing the purified paranemin IF network, indicating that it either is incorporating into heteropolymeric IFs. An analogous situation occurs with the two large neurofilament triplet proteins NF-H and NF-M, which only can assemble into IFs with the smaller triplet protein NF-L (74, 75). Our in vivo assembly studies with paranemin also are in concert with negative staining observations that paranemin by itself does not assemble into 10-nm diameter filaments when dialogized into IF-forming buffer. In summary, the sequence revealed several unexpected and significant properties of paranemin, including: 1) paranemin is an IF protein rather than an IFAP; 2) paranemin is a huge (178 kDa) IF protein, being much larger, for instance, than even neurofilament-H (~112 kDa), which is the largest of the triplet protein subunits composing neurofilaments; 3) paranemin shares several molecular features in common with the IF proteins nestin and tanabini, but significant differences as well; 4) paranemin contains approximately 36 consecutive, unusual pseudo-heptad repeats near the middle of the molecule, which have not previously been described in any of the over 60 IF proteins; and 5) regions within the complete sequence of paranemin are nearly identical to the partial sequences reported for two other proteins, EAP-300 and IFAPa-400. Transfection studies using SW13 cells have shown that expressed paranemin only is present in IFs in the presence of a major IF protein. Thus, paranemin likely functions within cells as a component of heteropolymeric IFs. As an IF component, paranemin should have an important role in modulating IF function in developing and mature cardiac and skeletal muscle cells and in other cells such as developing neurons. Taken in toto the results herein provide important information to help discover how the IFs in cells of muscle, nerves, and possibly other tissues, are linked to other cytoskeletal structures and regulate aspects of the cell cytoskeleton.

Acknowledgments—We thank Dr. Wallace Ip, University of Cincinnati School of Medicine, for providing the SW13 cell lines, Dr. Elisabeth Huff-Lonergan, Auburn University, for advice regarding immunolocalization studies, and assistance in the expression of the full-length paranemin cDNA, respectively.

REFERENCES

1. Steinitz, P. M., and Roop, D. R. (1988) Annu. Rev. Biochem. 57, 593–625
2. Robson, R. M. (1989) Curr. Opin. Cell Biol. 1, 36–42
3. Fuchs, E., and Weber, K. (1984) Annu. Rev. Biochem. 53, 345–382
4. Klymkowsky, M. W. (1995) Curr. Opin. Cell Biol. 7, 46–54
5. Fuchs, E. (1996) Annu. Rev. Genet. 30, 197–231
6. Fasani, F. (1997) BioEssays 19, 297–305
7. Lazarides, E. (1980) Nature 283, 249–256
8. Goldman, R. D., Khour, S., Chou, Y. H., Opal, P., and Steinitz, P. M. (1996) J. Cell Biol. 134, 971–983
9. Chou, Y. H., Skalli, O., and Goldman, R. D. (1997) Curr. Opin. Cell Biol. 9, 49–53
10. Maniotis, A. J., Chen, C. S., and Ingber, D. E. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 849–854
11. Huiatt, T. W., Robson, R. M., Arakawa, N., and Stromer, M. H. (1980) J. Biol. Chem. 255, 6981–6989
12. O'Shea, J. M., Robson, R. M., Hartzler, M. K., Huiatt, T. W., Rathbun, W. E., and Stromer, M. H. (1981) Biochem. J. 195, 345–356
13. Richardson, F. L., Stromer, M. H., Huiatt, T. W., and Robson, R. M. (1981) Eur. J. Cell Biol. 26, 91–100
14. Takano, E., Maki, M., Mori, H., Hatanaka, M., Marti, T., Kannagi, R., Ooi, T., and Murachi, T. (1988) Biochemistry 27, 1964–1972
15. Marston, S., and Smith, C. W. (1985) J. Biol. Chem. 260, 393–404
16. Chou, P. Y., and Fauman, E. D. (1978) Adv. Enzymol. 47, 45–148
17. Garnier, J., Oguthorpe, D. J., and Robson, B. (1978) J. Mol. Biol. 120, 97–120
18. Altucci, S. F., Gish, W., Miller, W., Meyers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 493–521
19. Frohman, M. A., Bush, M. K., and Martin, G. R. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8998–9002
20. Goll, D. E., Stromer, M. H., Robson, R. M., Luke, B. M., and Hammond, K. S. (1977) in Insect Flight Muscle (Tregear, R. T., ed) pp. 15–40, Elsevier/North-Holland Inc., New York
21. Sarria, A. J., Lieber, J. G., Nordeen, S. K., and Evans, R. M. (1994) J. Cell Sci. 107, 1593–1607
22. McCabe, C. F., Thompson, R. P., and Cole, G. J. (1992) Dev. Brain Res. 66, 11–23
23. Vincent, M., and Lahaie, C. (1988) Biochem. Cell Biol. 66, 184–192
24. Bryan, J. (1989) J. Muscle Res. Cell Motil. 10, 95–96
25. Graceffa, P., Wang, C.-L. A., and Stafford, W. F. (1988) J. Biol. Chem. 263, 14196–14202
26. Takeo, K., Maki, M., Mori, H., Hatanaka, M., Marti, T., Kannagi, R., Ooi, T., and Murachi, T. (1988) Biochemistry 27, 1964–1972
Characterization of the Novel IF Protein Paranemin