Molecular Characterization of a Novel Intracellular Hyaluronan-binding Protein*

Received for publication, March 31, 2000, and in revised form, June 6, 2000
Published, JBC Papers in Press, July 7, 2000, DOI 10.1074/jbc.M002737200

Lei Huang, Nicholas Grammatikakis, Masahiko Yoneda, Shib D. Banerjee, and Bryan P. Toole‡
From the Department of Anatomy and Cellular Biology, Tufts University School of Medicine, Boston, Massachusetts 02111

Hyaluronan has well defined functions in extracellular matrices and at the surface of cells. However, several studies have now shown that significant pools of hyaluronan are also present intracellularly, but its function therein is unknown. One avenue of investigation that may assist in defining the function of intracellular hyaluronan is to identify intracellular hyaluronan-binding proteins. In previous studies we identified CDC37, a cell cycle regulatory protein, using a monoclonal antibody that recognizes a novel group of hyaluronan-binding proteins. In this study, we have identified a second hyaluronan-binding protein with this antibody and characterized its properties. This protein, which we have termed IHABP4, was also found to be an intracellular and a specific hyaluronan-binding protein, containing several hyaluronan-binding motifs: (R/K)X7(R/K) (where R/K denotes arginine or lysine and X denotes non-acidic amino acids). Furthermore, we have determined the gene organization of IHABP4 and cloned cDNAs for the chick, mouse, and human homologs. Comparison of the deduced chick, mouse, and human protein sequences showed that the hyaluronan-binding motifs, (R/K)X7(R/K), in these sequences are conserved; both chick and mouse IHABP4 were shown directly to bind hyaluronan. Biochemical fractionation and immunofluorescent localization of epitope-tagged IHABP4 indicated that it is mainly present in the cytoplasm. These data support the possibility that intracellular hyaluronan and its binding proteins may play important roles in cell behavior.

Hyaluronan is a ubiquitous glycosaminoglycan (GAG) that is present in extracellular and pericellular matrices. On the basis of its unique physical and chemical properties, it has been well documented that hyaluronan plays a critical role in dynamic structural changes within extracellular matrices during development and tissue remodeling, as well as in maintenance of mechanical properties and homeostasis of many tissues (1, 2). Interactions of hyaluronan with hyaluronan-binding proteins (HABPs) are involved in establishing and modifying the structural properties of extracellular matrices. Furthermore, it is increasingly appreciated that hyaluronan is a crucial pericellular and cell surface component that actively participates in regulating cell behavior through its interactions with cell surface HABPs such as CD44 (1–4). A surprising new development, however, has been the discovery of several intracellular HABPs (IHABPs), but the interactions and functions of intracellular hyaluronan and IHABPs are not yet understood.

Convincing studies showing the existence of intracellular pools of hyaluronan in vivo and in vitro (5–8) and the discovery of several IHABPs, including CDC37 (9), RHAMM/IHABP (10, 11), and P32 (12), have raised an interesting question. Does intracellular hyaluronan, like extracellular hyaluronan, play an important role in regulating cellular behavior through interactions with IHABPs? Recently, it was reported that rapid uptake of Texas red-labeled hyaluronan and its subsequent accumulation in the cell processes, perinuclear area, and nucleus of transformed cells were associated with enhanced cell motility (13). Furthermore, careful examination confirmed the intracellular localization of endogenous hyaluronan as well as hyaluronan-binding sites and demonstrated dynamic changes in their pattern of distribution during proliferation of smooth muscle cells and fibroblasts (8). These data suggest strongly that intracellular hyaluronan has important roles in cellular processes.

Using the monoclonal antibody (mAb), IVd4, which recognizes a novel group of chick HABPs (14), we have previously identified and characterized CDC37, a cell cycle regulatory protein (9, 15). In the present study, we have characterized a cDNA, LH21, that encodes a second novel chick HABP recognized by mAb IVd4, using similar strategies as described previously (9, 15). This HABP, like CDC37, is also an intracellular protein containing multiple hyaluronan binding motifs, (R/K)X7(R/K), in which B denotes arginine or lysine and X denotes any amino acid except glutamic or aspartic acid (16). Furthermore, we have cloned the mouse and human homologs of this gene; shown that the hyaluronan binding motifs are conserved among the deduced chick, mouse, and human protein sequences; and demonstrated that the chick and mouse proteins bind hyaluronan specifically. These results imply that the hyaluronan-binding properties of this protein have functional importance. We have termed this protein IHABP4, because three other IHABPs have so far been characterized: CDC37 (9), RHAMM/IHABP (10, 11), and P32 (12).

EXPERIMENTAL PROCEDURES

Expression Library Screening with mAb IVd4—A chick embryo heart cDNA library, supplied by Drs. R. Markwald and E. Krug, University of S. Carolina Medical School, was screened with mAb IVd4 as described previously (9). One of the positive clones, which was repeatedly selected
and during independent screenings, is defined as LH21 for further characterization.

**3’- and 5’-RACE Reactions and Cloning of Mammalian Homologs—**

For cloning the full-length cDNA of chick LH21, a library ofadapter-digested double strand cDNA was made with poly(A)− RNA from day 6 chick limbs according to the manufacturer’s instructions (CLONTECH, Palo Alto, CA). The library was used as a template for performing both 5’- and 3’-RACE reactions based on the partial cDNA sequence of the initial clone LH21. The primer, F1, used for 3’-RACE was 5′-TTTCTACAGCTGCTGGACGCAGAGTTCTGAGGTTG-3′, corresponding to bases 245-267 of the cDNA sequence; and the primer, R1, used for 5’-RACE was 5′-CTGACATGGAAGAAAGGAAATAGGTTCCTCC-3′. PCR reactions were done under standard high stringency conditions. The primer pairs used for these two partial mammalian cDNA sequences, a similar strategy to that used for cloning full-length chick cDNA was used to clone full-length cDNAs of the mammalian homologs of LH21, with some modifications. Briefly, mouse and human heart poly(A)− RNAs were purchased from CLONTECH Laboratories, Inc. The cDNA templates for RACE reactions were made using the SMART RACE cDNA amplification kit (CLONTECH). The cDNA templates for PCR reactions were used as template. The resulting PCR products were first cloned into the PCR cloning vector PCRII to confirm their sequence authenticity through sequence analysis from both directions of the insert. After sequence confirmation, all inserts were cut out of the PCRII vector and subcloned into a mammalian expression vector, PCI-neo (Promega). Two reverse primers were designed to produce a common 27-bp sequence encoding a 9-amino acid hemagglutinin tag, followed by the specific chick LH21 cDNA sequences. Thus all resulting constructs could be expressed as recombinant proteins in mammalian expression kits containing the hemagglutinin tag. Four constructs were made: wtChLH21-Hatag, M1ChLH21-Hatag, M2ChLH21-Hatag, and M3ChLH21-Hatag. Primers used for making these chick LH21 constructs were as follows:

ChF1 primer was 5′-TCGGACCAATGATTAGAGGAGTGCTC-3′; ChF2, 5′-ATGACTGTCTCAGCTGACCTCC-3′; ChR1, 5′-TCAGGG- TATGAGGAGGAGGATGCTC-3′; ChR2, 5′-TAAGGCCGTTATGAGGAGGAGGATGCTC-3′. 

**Northern Blot Analysis—**

Hybridization and washes were determined by the double stranded DNA/dideoxy chain termination method using a Sequenase 2.0 kit (U.S. Biomedical Corp.). After RACE products were amplified by the SMART RACE cDNA amplification kit (CLONTECH), Northern RNA preparations were subjected to -RACE reactions, a GC-rich PCR system was used (CLONTECH).

**Sequence and Sequencing Analysis—**

For RACE products were cloned into the PCRII vector, the nucleotide sequences of these products were determined by the double stranded DNA/dideoxy chain termination method using a Sequenase 2.0 kit (U.S. Biomedical Corp.).

**Glycosaminoglycan-binding Assays—**

We used two assays to assess hyaluronan binding activity. The first was a [H]hyaluronan overlay assay as described previously (9, 14). Briefly, after separation by SDS-PAGE, proteins were transferred to a nitrocellulose membrane. The blot was incubated in a solution of 1.2 × 10−5 cpm [H]hyaluronan/mg of protein. As a negative control, water was added instead of [H]hyaluronan. The radioactivity bound to the blot was assayed using a liquid scintillation spectrometer.

**Construction and Expression of Epitope-tagged Recombinant Proteins—**

The complete coding region sequences of chick LH21 and three truncated mutant cDNAs (M1, M2, and M3; see Fig. 7A) were amplified by PCR reactions using different combinations of reverse and forward primer pairs (see below) with the cloned full-length chick LH21 cDNA as template. The resulting PCR products were first cloned into the PCR cloning vector PCRII to confirm their sequence authenticity through sequence analysis from both directions of the insert. After sequence

## New Intracellular Hyaluronan-binding Protein

Novel Intracellular Hyaluronan-binding Protein

During independent screenings, it was identified as LH21 for further characterization.

**3’- and 5’-RACE Reactions and Cloning of Mammalian Homologs—**

For cloning the full-length cDNA of chick LH21, a library of adapter-digested double strand cDNA was made with poly(A)− RNA from day 6 chick limbs according to the manufacturer’s instructions (CLONTECH, Palo Alto, CA). The library was used as a template for performing both 5’- and 3’-RACE reactions based on the partial cDNA sequence of the initial clone LH21. The primer, F1, used for 3’-RACE was 5′-TTTCTACAGCTGCTGGACGCAGAGTTCTGAGGTTG-3′, corresponding to bases 245-267 of the cDNA sequence; and the primer, R1, used for 5’-RACE was 5′-CTGACATGGAAGAAAGGAAATAGGTTCCTCC-3′, complementary to bases 345-322 of the sequence. Unique bands obtained from 3’- and 5’-RACE reactions were purified and ligated into the PCRII vector (Invitrogen) for sequence analysis. The full-length cDNA was then obtained by fusing the 3’- and 5’-RACE products by PCR using the Marathon kit according to the manufacturer’s instructions, followed by sequencing. The sequence was verified by direct amplification of the whole cDNA sequence by reverse transcription-PCR using a primer pair at the extreme 5’- and 3’-ends of the extended cDNA sequence with the double strand cDNA library as template.

For cloning the mammalian homologs of LH21, mouse and human EST as well as GenBank data bases were searched using the chick sequence with BLAST (Basic Local Alignment Sequence Tool) search programs. A set of partial mammalian cDNA sequence encoding a protein called Ki-1/5 antigen (GenBank accession number u77237) and a mouse EST sequence (accession number s29443) were found to be most likely candidates for the mammalian homologs of chick LH21. Based on these two partial mammalian cDNA sequences, a similar strategy to that used for cloning full-length chick cDNA was used to clone full-length cDNAs of the mammalian homologs of LH21, with some modifications. Briefly, mouse and human heart poly(A)− RNAs were purchased from CLONTECH Laboratories, Inc. The cDNA templates for RACE reactions were made using the SMART RACE cDNA amplification kit (CLONTECH). For 5’-RACE reactions, a GC-rich PCR system was used (CLONTECH).

**Sequence and Sequencing Analysis—**

For RACE products were cloned into the PCRII vector, the nucleotide sequences of these products were determined by the double stranded DNA/dideoxy chain terminations method using a Sequenase 2.0 kit (U.S. Biomedical Corp.).

**Northern Blot Analysis—**

For chick RNA preparation, whole chick embryos as well as chick embryonic tissues (limbs and brain) were used. Total poly(A)− RNA were prepared as described previously (15). Samples of the poly(A)− RNA preparations were subjected to electrophoresis in a formaldehyde gel and then transferred to a nitrocellulose membrane. The membrane was probed with chick clone LH21 cDNA labeled with [32P]dCTP using a random priming DNA labeling kit (Roche Molecular Biochemicals). Hybridization and washes were done under standard high stringency conditions.

For Northern blot analysis of mouse mRNAs, a nitrocellulose membrane preblotted with multiple poly(A)− RNAs prepared from different mouse tissues was purchased from CLONTECH. Probe preparation, hybridization, and controls were conducted according to the manufacturer’s instructions.

**Determination of Gene Organization by PCR—**

The genomic region encompassing the chick LH21 cDNA sequence was obtained by amplifying several overlapping genomic fragments, because the size of this genomic region is too large for direct amplification (>20 kb). Cloning of PCR products and identification of boundary sequences between introns and exons were determined by the same strategy as previously used (15).

**Construction and Expression of Epitope-tagged Recombinant Proteins—**

The complete coding region sequences of chick LH21 and three truncated mutant cDNAs (M1, M2, and M3; see Fig. 7A) were amplified by PCR reactions using different combinations of reverse and forward primer pairs (see below) with the cloned full-length chick LH21 cDNA as template. The resulting PCR products were first cloned into the PCR cloning vector PCRII to confirm their sequence authenticity through sequence analysis from both directions of the insert. After sequence confirmation, all inserts were cut out of the PCRII vector and subcloned into a mammalian expression vector, PCI-neo (Promega). Two reverse primers were designed to produce a common 27-bp sequence encoding a 9-amino acid hemagglutinin tag, followed by the specific chick LH21 cDNA sequences. Thus all resulting constructs could be expressed as recombinant proteins in mammalian expression kits containing the hemagglutinin tag. Four constructs were made: wtChLH21-Hatag, M1ChLH21-Hatag, M2ChLH21-Hatag, and M3ChLH21-Hatag. Primers used for making these chick LH21 constructs were as follows:

ChF1 primer was 5′-TCGGACCAATGATTAGAGGAGTGCTC-3′; ChF2, 5′-ATGACTGTCTCAGCTGACCTCC-3′; ChR1, 5′-TCAGGG- TATGAGGAGGAGGATGCTC-3′; ChR2, 5′-TAAGGCCGTTATGAGGAGGAGGATGCTC-3′.
solution containing 1% CPC and 30 mM NaCl for each wash, and finally suspended in 50 mM NaCl. Samples were then subjected to SDS-PAGE and Western blot analysis with rat monoclonal antibody (3F10 clone, Roche Molecular Biochemicals) against hemagglutinin tag as described above. Untreated lysate was loaded as a positive control. The relative densities of the bands obtained in the Western blots were determined using a BioImage whole-band analysis package (Millipore).

**Immunofluorescent Staining—**NIH3T3 cells, cultured in 4-well chamber slides (Falcon) in Dulbecco’s modified Eagle’s medium plus 10% fetal bovine serum (Hyclone), were transfected with pure plasmid DNA for either amino- or carboxyl-hemagglutinin-tagged cDNA constructs. At 48 h post-transfection, cells were fixed with 4% formaldehyde for 10 min at room temperature then permeabilized with 0.3% Triton X-100 in PBS for 10 min. The expressed protein was detected by rat monoclonal antibody (3F10 clone, Roche Molecular Biochemicals) against hemagglutinin tag followed by secondary antibody conjugated with rhodamine (Pierce). Controls for protein staining were non-permeabilized cells, preincubation of primary antibody with tag peptide, or incubation with rhodamine-conjugated secondary antibody alone.

**RESULTS**

**Cloning of Chick LH21 cDNA**—In previous studies (9, 15), we reported the molecular characterization of a chick IHABP, CDC37. The cDNA for CDC37 was isolated from a chick embryo heart cDNA library by immunoscreening with mAb IVd4. This antibody was selected by screening hybridomas raised against a partially purified preparation of chick embryo HABPs in the presence and absence of polymeric and oligomeric hyaluronan. It recognizes a novel group of HABPs present in a variety of tissue and cell types, especially in the embryo (14). Here we report characterization of a cDNA, LH21, for a second IHABP recognized by mAb IVd4, which we have termed IHABP4.

We chose to investigate the cDNA clone LH21, because, like the original clone for CDC37, it displayed consistently strong immunoreactivity with mAb IVd4 during successive rounds of immunoscreening of the chick heart cDNA expression library. The LH21 clone was isolated and sequenced by the same strategy as described previously for CDC37 (9). The size of the cDNA insert was about 1.2 kb, but initial examination of the sequence suggested that the insert might be a partial cDNA sequence. Northern blot analysis was performed using labeled LH21 cDNA to hybridize with mRNA preparations from chick embryo HABPs in the presence and absence of polymeric and oligomeric hyaluronan. The result (Fig. 1B) showed a prominent band slightly larger in size than 18 S ribosomal RNA, and a weak band similar in size to 27 S ribosomal RNA. Because mRNA was used rather than total RNA, the signals were not due to ribosomal RNAs. Furthermore, in a separate Northern blot analysis, total RNAs prepared from chick whole embryo, brain, and limb were loaded in the same gel together with the mRNA preparations. Strong signals were detectable only from the mRNA samples, not from the total RNA samples (data not shown), indicating that the transcription level of LH21 in the embryo is relatively low, and confirming that the signals from the mRNA preparations in Fig. 1B were specific. By comparing the size of the LH21 transcript (∼2.0 kb as estimated from Fig. 1B) with that of the LH21 cDNA we had cloned (∼1.2 kb), it was clear that clone LH21 was incomplete.

To extend the sequence of cDNA LH21, we designed primers F1 and R1 (Fig. 1A) for 3′- and 5′-RACE reactions, respectively, according to the known partial sequence of the clone. At least 10 independent clones for each PCR product from 3′- and 5′-RACE reactions were sequenced. By fusing the 3′- and 5′-RACE products through a PCR reaction, we now obtained an extended LH21 cDNA of ∼1.8 kb. To eliminate any possible artifact during PCR and cloning, the cloned cDNA sequence was further confirmed by direct amplification of the whole cDNA by reverse transcription-PCR, using a chick embryo limb cDNA library as template. The size and sequence of this di...
directly amplified PCR product were the same as that of the fused cDNA. The size of the extended LH21 cDNA (~1.8 kb) plus a poly(A) tail is consistent with that of the major band from Northern blot analysis (~2.0 kb, in Fig. 1B). The putative translational start codon conforms to the Kozak consensus sequence around a translation initiation site; furthermore, the presence of an in-frame stop codon, upstream of the start codon in the 5’-untranslated region, supports the prediction of an open reading frame. At the end of the 3’-UTR is a typical polyadenylation signal, AATAAA. Thus, we believe that this cDNA represents the full-length LH21 cDNA for chick IHABP4. The GenBank accession number for the chick sequence is AF227683.

The extended cDNA sequence described above contains an open reading frame of 1071 bases, encoding a putative 357-amino acid polypeptide with an expected molecular mass of 42 kDa (see Fig. 7B, lane 1), consistent with the size of cDNA.

Examination of the deduced amino acid sequence of IHABP4 (Fig. 1A) revealed multiple hyaluronan binding motifs: -(R/K)X7(R/K)- (16). These motifs are between amino acid residues Lys-51 to Arg-59, Arg-151 to Lys-173 (which contains three overlapped -(R/K)X7(R/K)- motifs), and Lys-267 to Lys-275.
ately downstream from 3'-end of an exon. The organization of the \( \text{IHABP4} \) gene is shown in Fig. 2. The size of the genomic region covering the \( \text{LH21} \) cDNA is about 20 kb. There are at least seven exons and six introns in the \( \text{IHABP4} \) gene, and all boundaries between introns and exons comply with the GT (5'-end of the intron) and AG (3'-end of the intron) rules for splicing sites (Fig. 2B). The size of each intron was determined by PCR as described previously and found to range from ~1.0 to ~6.0 kb.

**Cloning of Mouse and Human \( \text{IHABP4} \) Homologs**—To determine whether any possible homologs of chick \( \text{IHABP4} \) were present in the data bases, a similarity search was performed using BLAST programs. Significant similarity to the deduced \( \text{LH21} \)-encoded protein sequence was obtained for a partially characterized human protein of unknown function, called \( \text{Ki-1/57} \) (21). \( \text{Ki-1/57} \) is an intracellular protein recognized by the \( \text{Ki-1} \) monoclonal antibody originally developed against an apparently unrelated cell surface protein, CD30 (22). However, the available cDNA sequence encodes only a partial protein sequence; the sequence encoding the amino-terminal part of the protein is missing due to difficulties encountered in cloning the full-length cDNA (21). In addition, we found several newly released mouse EST sequences, which have very high similarity to human \( \text{Ki-1/57} \) cDNA. Based on the above sequence information, we assumed that the cDNA sequence of human \( \text{Ki-1/57} \) and the mouse EST sequences could represent portions of cDNAs for mammalian homologs of chick \( \text{IHABP4} \), and we began attempts to clone the full-length cDNA sequences for both human and mouse \( \text{IHABP4} \).

We first assembled selected mouse EST sequences to obtain a maximum partial sequence, and then confirmed the assembled sequence by direct reverse transcription-PCR using mouse heart mRNA as template. Because the nature of the 5'-part of the mouse cDNA might be very similar to that of the human cDNA and because rigid GC-rich regions are often found at the 5'-region of mRNAs, we introduced a commercially available GC-rich PCR system into our 5'-RACE reactions. This strategy enabled us to obtain the 5'-part of mouse cDNA (Fig. 3A). As we predicted, the cloned 5'-region (1400 bp from the 3'-end) of mouse cDNA shows very high GC content (~75%, see Fig. 3A). Interestingly, multiple deletions reducing the GC content and overall size of this region were found in the corresponding part of chick \( \text{LH21} \) cDNA (Figs. 1A and 4), thus explaining the lack of problems encountered during chick \( \text{LH21} \) cloning. As a result, the 3'-part of mouse cDNA was successfully obtained from 3'-RACE reactions, but the 5'-RACE reactions did not give any distinct band even after numerous trials with different primers, high temperature RT reactions, and hot-start \( \text{Taq} \) DNA polymerase enzymes. Similarly, the authors who partially cloned human \( \text{Ki-1/57} \) cDNA could not obtain the 5'-part of the human cDNA and suggested that a rigid secondary structure at the 5'-end of the transcript might prevent reverse transcription by conventional methods (21). Because the nature of the 5'-part of the mouse cDNA might be very similar to that of the human cDNA and because rigid GC-rich regions are often found at the 5'-region of mRNAs, we introduced a commercially available GC-rich PCR system into our 5'-RACE reactions. This strategy enabled us to obtain the 5'-part of mouse cDNA (Fig. 3A). As we predicted, the cloned 5'-region (~400 bp from the 5'-end) of mouse cDNA shows very high GC content (>75%, see Fig. 3A). Interestingly, multiple deletions reducing the GC content and overall size of this region were found in the corresponding part of chick \( \text{LH21} \) cDNA (Figs. 1A and 4), thus explaining the lack of problems encountered during chick \( \text{LH21} \) cDNA cloning. The cloned mouse cDNA is about 2.5 kb in size and contains a 1233-base open reading frame region encoding 411 amino acids (Fig. 3A).

**Fig. 3. Characterization of a cDNA for mouse \( \text{IHABP4} \).** A, nucleotide and deduced amino acid sequences. Nucleotides are numbered on the left and amino acids on the right. The start and stop codons flanking the open reading frame are shown in boldface. An in-frame stop codon within 5'-UTR is underlined. The hyaluronan-binding motifs are underlined and boldface. B, Western blot analysis of transiently expressed mouse \( \text{IHABP4} \) using monoclonal antibody against hemagglutinin tag. Lanes 1, 2, and 3 are the cell lysates extracted from CHO-K1 cells at 48-h post-transfection with vector alone and C-tagged and N-tagged mouse \( \text{IHABP4} \) constructs, respectively.
To confirm that the deduced mouse IHABP4 protein is encoded by the cloned cDNA, we made two epitope-tagged mouse IHABP4 constructs. One of these is tagged with hemagglutinin at the amino terminus and the other at the carboxyl terminus so as to eliminate possible interference with expression, hyaluronan binding, or immunostaining by the 9-amino acid hemagglutinin tag at either end of the protein. Transient expression of mouse IHABP4 in CHO-K1 cells showed that the recombinant protein was equally expressed from either the N-tagged or C-tagged construct, and the size of the protein in each case was estimated to be 55 kDa (Fig. 3B). Thus the mouse IHABP4 protein is indeed encoded by the cloned cDNA. The molecular mass of mouse IHABP4 is in good agreement with the molecular mass of human Ki-1/57 (57 kDa) reported previously (21).

We also completed the missing 5′ part of the partial cDNA sequence of human Ki-1/57 by the same strategy used for cloning mouse cDNA. The deduced peptide sequence was extended from an original 299 amino acids to 413 (Fig. 4). Because extremely high similarity (~90% identity) was found between mouse IHABP4 protein and the completed human Ki-1/57 antigen (Fig. 4), we regard the Ki-1/57 as the human IHABP4 homolog. The GenBank accession number for the human sequence is AF241831.

In the deduced mouse IHABP4 protein, as well as in the extended human Ki-1/57 protein, the hyaluronan-binding motifs, -(R/K)X7(R/K)-, were found to be well conserved in corresponding regions to those in chick (Fig. 4). This supported the possibility that IHABP4 may be a novel hyaluronan-binding protein. Comparison of chick IHABP4 with its mammalian homologs (Fig. 4) reveals ~48% identity and ~60% similarity between chick and mouse or chick and human proteins. The carboxyl-terminal region (~130 amino acids) is more conserved among all three species (~80% identity) than the rest of the protein and also contains a virtually identical -(R/K)X7(R/K)- motif, KAVVHKS(K/R), suggesting that this region might be a critical region for the function of this protein.

We conducted further BLAST searches using the mouse and human sequences. In addition to Ki-1/57, another human protein, CGI-55 (GenBank accession numbers AAD34050 and NP056455), was found to have a high degree of homology to IHABP4 but is of unknown function. Partial homology was also

FIG. 4. Comparison of chick IHABP4 with mammalian homologs. Identical amino acid residues at the same positions in the sequences are marked with black boxes; conserved residues are marked with shadowed boxes. Dots indicate gaps or deletions in the sequences. Hyaluronan-binding motifs are underlined. Note: the hyaluronan-binding motif near the amino termini of the three species is marked with a dashed line. In this case, the chick motif (residues Lys49 to Arg58) is not strictly aligned with the human and mouse motifs (residues Arg49 to Arg58). GenBank accession numbers for these sequences are AF227683 (chick), AF227684 (mouse), and AF241831 (human).

To confirm that the deduced mouse IHABP4 protein is encoded by the cloned cDNA, we made two epitope-tagged mouse IHABP4 constructs. One of these is tagged with hemagglutinin at the amino terminus and the other at the carboxyl terminus so as to eliminate possible interference with expression, hyaluronan binding, or immunostaining by the 9-amino acid hemagglutinin tag at either end of the protein. Transient expression of mouse IHABP4 in CHO-K1 cells showed that the recombinant protein was equally expressed from either the N-tagged or C-tagged construct, and the size of the protein in each case was estimated to be ~55 kDa (Fig. 3B). Thus the mouse IHABP4 protein is indeed encoded by the cloned cDNA. The molecular mass of mouse IHABP4 is in good agreement with the molecular mass of human Ki-1/57 (57 kDa) reported previously (21).

We also completed the missing 5′ part of the partial cDNA sequence of human Ki-1/57 by the same strategy used for cloning mouse cDNA. The deduced peptide sequence was extended from an original 299 amino acids to 413 (Fig. 4). Because extremely high similarity (~90% identity) was found between mouse IHABP4 protein and the completed human Ki-1/57 antigen (Fig. 4), we regard the Ki-1/57 as the human IHABP4 homolog. The GenBank accession number for the human sequence is AF241831.

In the deduced mouse IHABP4 protein, as well as in the extended human Ki-1/57 protein, the hyaluronan-binding motifs, -(R/K)X7(R/K)-, were found to be well conserved in corresponding regions to those in chick (Fig. 4). This supported the possibility that IHABP4 may be a novel hyaluronan-binding protein. Comparison of chick IHABP4 with its mammalian homologs (Fig. 4) reveals ~48% identity and ~60% similarity between chick and mouse or chick and human proteins. The carboxyl-terminal region (~130 amino acids) is more conserved among all three species (~80% identity) than the rest of the protein and also contains a virtually identical -(R/K)X7(R/K)- motif, KAVVHKS(K/R), suggesting that this region might be a critical region for the function of this protein.

We conducted further BLAST searches using the mouse and human sequences. In addition to Ki-1/57, another human pro-
noted with regions of some putative nucleic acid-binding proteins from invertebrates, e.g., the Drosophila gene products vig (accession number AAF44918) and CG11844 (accession number AAF56404). Although these proteins have some relationship to RNA-binding proteins, they do not contain RNA recognition domains as assessed using the SMART program. We also used the SMART program to search for RNA recognition domains in IHABP4 but found none. We used several programs to search for other possible motifs and simple modules in IHABP4 to obtain clues as to its subcellular localization and its structural relationship to other proteins, but no such motifs were revealed.

Expression of IHABP4 mRNA in Mouse Tissues—To investigate mouse IHABP4 expression pattern in mouse tissues, and to examine whether the cloned mouse cDNA is close in size to its transcript, Northern blot analysis was conducted. As shown in Fig. 5, mouse heart, brain, liver, and kidney tissues express a prominent transcript of ~2.5 kb, consistent with the size of cDNA we have cloned. The origin of a faint band at >5.0 kb is unclear. However, spleen, lung, and skeletal muscle tissues do not express IHABP4 mRNA, or express very low levels, indicating that IHABP4 is unlikely to be a general housekeeping gene. To our surprise, the major transcript of IHABP4 in testis is much more abundant and smaller (~1.35 kb) than that in other tissues. Two minor transcripts around 2.5 and 4.0 kb are also expressed in the testis. The reasons for these differences are not known, but alternative splicing, promoter usage, or poly(A) usage are possibilities.

Hyaluronan-binding Properties of IHABP4—As discussed above, the amino acid sequence of IHABP4 includes several conserved hyaluronan-binding motifs. However, because it is not yet known whether these motifs are in fact hyaluronan-binding sites within IHABP4, we further examined whether IHABP4 protein actually binds hyaluronan, using two approaches.

For the first procedure, we expressed the recombinant chick IHABP4 as a fusion protein in bacteria (9). The bacterial proteins were then separated by SDS-PAGE, transferred to nitrocellulose membrane, and cut into strips. Separate strips were overlaid with [3H]hyaluronan or the antibody IVd4, which recognizes IHABP4 (9, 14). A single band was obtained in the IVd4 Western blot, and this band corresponded with the position of [3H]hyaluronan binding (Fig. 6A). Addition of hyaluronan oligo-
Glycosaminoglycans inhibited binding of [3H]hyaluronan, indicating that this binding was specific. In addition, extracts of bacteria lacking IHABP were analyzed and found to exhibit no detectable binding of IVd4 or [3H]hyaluronan (data not shown).

In the second procedure, we produced stable transfectants expressing N- and C-tagged IHABP4. We then tested binding of both N-tagged and C-tagged mouse IHABP4 to various GAGs, using a solution assay in which cell lysates containing tagged IHABP4 are mixed with GAG, and then co-precipitation of IHABP4 with GAG by CPC is measured (20). As can be seen in Fig. 6B, N-tagged mouse IHABP4 was efficiently co-precipitated after preincubation with exogenous hyaluronan and addition of CPC. However, very small amounts of IHABP4 were co-precipitated with heparan sulfate or chondroitin sulfate. The amounts of co-precipitated IHABP4 obtained in the presence of the various reagents was measured by densitometry and expressed relative to the amount co-precipitated with hyaluronan. The percentages relative to hyaluronan were 0.9% with heparan sulfate, 4.0% with chondroitin sulfate, and zero with hyaluronan oligomers, indicating that the interaction between mouse IHABP4 and hyaluronan is specific and not simply due to charge interactions. We also performed the assay using C-tagged mouse IHABP4 and obtained the same result as above (data not shown). These results demonstrate that IHABP4 is a specific HABP.

Because IHABP4 may be related to putative nucleic acid-binding proteins, we also determined whether IHABP4 binds RNA. To do this we used the second procedure above, because RNA binds to CPC in a similar fashion to GAGs. Some binding of IHABP4 was obtained but to a much smaller degree (6.7%) than with hyaluronan (Fig. 6B).

**Intracellular Localization of IHABP4**—A 22-amino acid hydrophobic stretch that complies with the Von Heijne (23) rule for signal peptide cleavage is present at the amino terminus of the deduced IHABP4 protein. However, only very small amounts of IHABP4 were co-precipitated with heparan sulfate or chondroitin sulfate. The amounts of co-precipitated IHABP4 obtained in the presence of the various reagents was measured by densitometry and expressed relative to the amount co-precipitated with hyaluronan. The percentages relative to hyaluronan were 0.9% with heparan sulfate, 4.0% with chondroitin sulfate, and zero with hyaluronan oligomers, indicating that the interaction between mouse IHABP4 and hyaluronan is specific and not simply due to charge interactions. We also performed the assay using C-tagged mouse IHABP4 and obtained the same result as above (data not shown). These results demonstrate that IHABP4 is a specific HABP.

Because IHABP4 may be related to putative nucleic acid-binding proteins, we also determined whether IHABP4 binds RNA. To do this we used the second procedure above, because RNA binds to CPC in a similar fashion to GAGs. Some binding of IHABP4 was obtained but to a much smaller degree (6.7%) than with hyaluronan (Fig. 6B).

**Intracellular Localization of IHABP4**—A 22-amino acid hydrophobic stretch that complies with the Von Heijne (23) rule for signal peptide cleavage is present at the amino terminus of the deduced IHABP4 protein. However, only very small amounts of IHABP4 were co-precipitated with heparan sulfate or chondroitin sulfate. The amounts of co-precipitated IHABP4 obtained in the presence of the various reagents was measured by densitometry and expressed relative to the amount co-precipitated with hyaluronan. The percentages relative to hyaluronan were 0.9% with heparan sulfate, 4.0% with chondroitin sulfate, and zero with hyaluronan oligomers, indicating that the interaction between mouse IHABP4 and hyaluronan is specific and not simply due to charge interactions. We also performed the assay using C-tagged mouse IHABP4 and obtained the same result as above (data not shown). These results demonstrate that IHABP4 is a specific HABP.

Because IHABP4 may be related to putative nucleic acid-binding proteins, we also determined whether IHABP4 binds RNA. To do this we used the second procedure above, because RNA binds to CPC in a similar fashion to GAGs. Some binding of IHABP4 was obtained but to a much smaller degree (6.7%) than with hyaluronan (Fig. 6B).

**Intracellular Localization of IHABP4**—A 22-amino acid hydrophobic stretch that complies with the Von Heijne (23) rule for signal peptide cleavage is present at the amino terminus of the deduced IHABP4 protein. However, only very small amounts of IHABP4 were co-precipitated with heparan sulfate or chondroitin sulfate. The amounts of co-precipitated IHABP4 obtained in the presence of the various reagents was measured by densitometry and expressed relative to the amount co-precipitated with hyaluronan. The percentages relative to hyaluronan were 0.9% with heparan sulfate, 4.0% with chondroitin sulfate, and zero with hyaluronan oligomers, indicating that the interaction between mouse IHABP4 and hyaluronan is specific and not simply due to charge interactions. We also performed the assay using C-tagged mouse IHABP4 and obtained the same result as above (data not shown). These results demonstrate that IHABP4 is a specific HABP.

Because IHABP4 may be related to putative nucleic acid-binding proteins, we also determined whether IHABP4 binds RNA. To do this we used the second procedure above, because RNA binds to CPC in a similar fashion to GAGs. Some binding of IHABP4 was obtained but to a much smaller degree (6.7%) than with hyaluronan (Fig. 6B).
NIH 3T3 cells were transfected with N-tagged mouse control for background staining. The remaining cells provide an internal contrast image of the same field. Note that only some of the cells express IHABP4 (two in this field), because the cells are transiently transfected and not cloned. The majority of expressed protein was found in the cytosolic fraction; about 30% was found in membrane and other organelle fractions. We conclude that chick IHABP4 is an intracellular protein, mainly present in the cytoplasm. Stable transfectants of N- and C-tagged mouse constructs in CHO-K1 cells were also used for subcellular fractionation. Subcellular fractionation gave a similar distribution as found for chick IHABP4, i.e. the majority (~70%) of the IHABP4 protein was found in the cytosolic fraction and ~30% in other membrane and organelle fractions (data not shown).

We used immunocytochemistry to examine the distribution of IHABP4 in cells transiently transfected with hemagglutinin-tagged IHABP4. IHABP4 was visualized with antibody to hemagglutinin tag. IHABP4 staining was seen in the cytoplasm as a diffuse, network-like pattern, especially in the perinuclear region (Fig. 8). IHABP4 staining was completely eliminated in the presence of the peptide tag or by omitting primary antibody; also, no IHABP4 staining was observed in non-permeabilized cells (data not shown).

**DISCUSSION**

In this study, we have characterized a novel and specific IHABP. Until the function of this protein is better understood, we propose to call it IHABP4, because three other IHABPs have been convincingly characterized, namely CDC37 (9), RHAMM/IHABP (10, 11), and P32 (12). IHABP4 contains multiple hyaluronan-binding motifs throughout its sequence, notably one stretch with three overlapping motifs in the middle region of the protein sequence. These motifs are conserved in the chick, human, and mouse sequences, suggesting that IHABP4 may have a hyaluronan-related function. This conclusion is supported by our data showing that both chick and mouse IHABP4 bind hyaluronan in a specific fashion. However, regions within IHABP4 may be related to putative nucleic acid-binding proteins, e.g. the *Drosophila* ving product. Although IHABP4 does not contain an established RNA recognition domain, we tested IHABP4 for RNA-binding activity and found it to exhibit weak binding compared with hyaluronan. Nevertheless, until further functional work is carried out, the question of whether hyaluronan is the natural ligand for IHABP4 must remain open. The homologies found between IHABP4, the human protein CGI-55, and the above-mentioned *Drosophila* proteins suggest that IHABP4 belongs to a family of related genes.

Unlike the other known IHABPs referred to above, IHABP4 mRNA is not ubiquitously expressed in adult tissues. IHABP4 mRNA is found at significant levels in adult heart, brain, liver, kidney, and testis, as well as in embryonic tissues, but not in adult spleen, lung, or skeletal muscle. A very prominent transcript that is smaller than that found elsewhere was seen only in testis. It is not yet known whether a different protein product from that found in this study results from this smaller transcript. The human homolog, also known as Ki-1/57, is expressed in activated, but not resting, human lymphocytes and is enriched in some tumor cells (21). Obviously, to understand the function of this novel HABP, further investigation will be needed, including examination of the expression levels of its mRNA and protein in various normal cells versus corresponding tumor cells, as well as its interaction with other regulatory or structural protein(s). In reference to the latter, previous evidence suggests that Ki-1/57 is associated with intracellular kinase activity (21).

A considerable amount of evidence has shown that sulfated GAG chains and proteoglycans are present in the cytoplasm and nucleus of a variety of cell types and that in some cases these intracellular GAG populations are likely to be involved in regulation of cell behavior. For example, it has been shown that heparan sulfate can act as a trans-repressor that interferes with the action of c-Fos and c-Jun on transcription events in *vivo*, and preliminary evidence has suggested that endogenous nuclear heparan sulfate has such a regulatory function in *vivo* (25). In support of this possibility, specific inhibitory subpopulations of heparan sulfate are targeted to the nucleus during cell proliferation (26, 27). Also, dynamic changes occur in the levels of biglycan and glypican proteoglycans in the nucleus during the cell cycle, and the core proteins of these proteoglycans have been shown to contain nuclear localization motifs (28). There is also definitive evidence that heparin is essential for retention of several proteases present within secretory granules of mast cells (29).

It has become increasingly evident that hyaluronan is also
Various studies discussed above is the exact mechanism whereby intracellular hyaluronan might influence cellular events. It is possible that the high level of hydration associated with hyaluronan (2) also plays a role in structural changes in the cytoskeleton or nuclear matrix during cell division or motility, e.g. in regulating cell shape or volume changes. Interestingly, recent evidence suggests that hyaluronan-associated changes in hydration play an important role in growth plate expansion during bone development (44) but that most of the hyaluronan at this site is intracellular (45). Irrespective of the role of intracellular hyaluronan, it is to be expected that its functions will be regulated and/or mediated by IHaBP. However, direct evidence for intracellular interaction in situ and for the functional consequence of these interactions is needed to elucidate this possibility.

Acknowledgment—We thank Aliki Grammatikakis for excellent technical assistance.

REFERENCES
1. Toole, B. P. (1999) in Guidebook to the Extracellular Matrix, Anchor, and Adhesion Proteins (T. Kreis, T., and Vale, R., eds) 2nd Ed., pp. 430–433, Oxford University Press, Oxford.
2. Toole, B. P. (2000) in Proteoglycans: Structure, Biology and Molecular Interactions (Iozzo, R., ed) pp. 61–92, Marcel Dekker, NY.
3. Kincade, P. W., Zheng, Z., Katoch, S., and Hanson, L. (1997) Curr. Opin. Cell Biol., 9, 635–642.
4. Entwistle, J., Hall, C. L., and Turley, E. A. (1996) J. Cell. Biochem., 61, 569–577.
5. Fujikawa, K., and Terayama, H. (1979) Biochem. Biophys. Acta., 565, 575–588.
6. Ripellino, J. A., Baito, M., Margolis, R. U., and Margolis, R. K. (1998) J. Cell. Biol., 106, 845–855.
7. Eggli, P., and Graber, W. (1999) J. Histochem. Cytochem., 47, 689–697.
8. Evanko, S. P., and Wight, T. N. (1999) J. Histochem. Cytochem., 47, 1331–1342.
9. Grammatikakis, N., Grammatikakis, A., Neveda, M., Yu, Q., Banerjee, S. D., and Toole, B. P. (1995) J. Biol. Chem., 270, 16168–16205.
10. Entwistle, J., Hall, C. L., and Turley, E. A. (1996) J. Cell. Biochem., 61, 569–577.
11. Assmann, V., Marshall, J. F., Fieber, C., Hofmann, M., and Hart, I. R. (1998) J. Cell Sci., 111, 1685–1694.
12. Do, T. B., and Datta, K. (1996) J. Biol. Chem., 271, 2206–2212.
13. Collis, L., Hall, C., Lange, L., Zebell, M., Prestwich, R., and Turley, E. A. (1998) FEBS Lett., 440, 444–449.
14. Banerjee, S. D., and Toole, B. P. (1991) Dev. Biol., 146, 186–197.
15. Huang, L., Grammatikakis, N., and Toole, B. P. (1998) J. Biol. Chem., 273, 3598–3603.
16. Yang, B., Yang, B. L., Savani, R. C., and Turley, E. A. (1994) EMBO J., 13, 286–296.
17. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) Protein Eng., 10, 1–6.
18. Schultz, M., Miletz, F., Bork, P., and Ponting, C. P. (1998) Proc. Natl. Acad. Sci. U. S. A., 95, 5857–5864.
19. Schultz, J., Copley, R. R., Doers, T., Ponting, C. P., and Bork, P. (2000) Nucleic Acids Res., 28, 347–354.
20. Steen, J. P., Kondo, K., Moll, J., Ponta, H., and Herrlich, P. (1997) J. Biol. Chem., 272, 31837–31844.
21. Kobarg, J., Schnittger, S., Fonatsch, C., Lemke, H., Bowen, M. A., Bock, F., and Hansen, H. P. (1997) Exp. Clin. Immunogenet., 14, 273–280.
22. Schwab, U., Hein, S., Sperling, J. C., Kirchner, H., Schaadt, M., and Dohl, V. (1992) Nature, 309, 65–67.
23. Von Heijne, G. (1994) J. Mol. Biol., 173, 243–251.
24. Herber, L. D., Kodukula, K., and Udenfriend, S. (1992) J. Biol. Chem., 267, 12168–12173.
25. Buesch, S. J., Martin, G. A., Barnhart, R. L., Mano, M., Cardin, A., and Jackson, R. L. (1992) J. Cell Biol., 116, 31–42.
26. Fedarko, N. S., and Conrad, H. E. (1986) J. Cell. Biol., 102, 587–589.
27. Fedarko, N. S., Ishihara, M., and Conrad, H. E. (1989) J. Cell. Physiol., 139, 267–294.
28. Liang, Y., Haring, M., Roughley, P. J., Margolis, R. K., and Margolis, R. U. (1997) J. Cell. Biol., 139, 851–864.
29. Humphries, D. E., Wong, K. W., Friend, D. S., Gush, M. F., Qiu, W. T., Huang, C., Sharpe, A. H., and Stevens, R. L. (1999) Nature, 400, 769–772.
30. Evanko, S. P., Angello, J. C., and Wight, T. N. (1999) Arterioscler. Thromb. Vasc. Biol., 19, 1004–1013.
31. Brechbi, M., Mayer, U., Schlosser, E., and Prehm, P. (1986) Biochem. J., 239, 445–450.
32. Rode, D., Hansen, H., Hafner, M., Lange, H., Mielle, V., Hansmann, M. L., and Lemke, H. (1992) J. Cell. Pathol., 140, 473–482.
33. Kosaki, R., Watanabe, K., and Yamaguchi, Y. (1999) Cancer Res., 59, 1141–1145.
34. Peterson, R. M., Yu, Q., Stamenkovic, I., and Toole, B. P. (2000) Am. J. Pathol., 156, 2159–2167.
35. Itano, N., Sawai, T., Miyaiishi, O., and Kimata, K. (1999) Cancer Res., 59, 2499–2504.
36. Yu, Q., Toole, B. P., and Stamenkovic, I. (1997) J. Exp. Med., 186, 1985–1996.
37. Yu, Q., and Stamenkovic, I. (1999) Genes Dev., 13, 35–48.
38. Culyt, M., Shizari, M., Thompson, E. W., and Underhill, C. B. (1994) J. Cell. Physiol., 160, 275–284.
39. Liu, D., Pearlman, E., Diaconu, E., Guo, K., Mori, H., Huqqi, T., Markowitz, S.,
    Willson, J., and Sy, M. S. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 7832–7837
40. Olofsson, A. M., Vestberg, M., Herwald, H., Rygaard, J., David, G., Arfors,
    K. E., Linde, V., Flodgaard, H., Dedio, J., Muller-Esterl, W., and Lundgren-
    Akerlund, E. (1999) J. Clin. Invest. 104, 885–894
41. Lord, J. M., and Roberts, L. M. (1998) J. Cell Biol. 140, 733–736
42. Cleves, A. (1997) Curr. Biol. 7, R318–R320
43. Weigel, P. H., Hascall, V. C., and Tammi, M. (1997) J. Biol. Chem. 272, 13997–14000
44. Pavasant, P., Shizari, T., and Underhill, C. B. (1996) J. Cell Sci. 109, 327–334
45. Takada, Y., Sakiyama, H., Kuriwai, K., Masuda, R., Inoue, N., Nakagawa, K.,
    Itano, N., Saito, T., Yamada, T., and Kimata, K. (1999) Cell Tissue Res. 298, 317–325