APBP-1, a DNA/RNA-binding Protein, Interacts with the Chick Aggrecan Regulatory Region*

Received for publication, May 17, 2005, and in revised form, July 19, 2005 Published, JBC Papers in Press, August 12, 2005, DOI 10.1074/jbc.M505380200

Edward W. Pirok III†§, Miriam S. Domowicz†, Judith Henry†, Youli Wang†, Matthew Santore†, Melissa M. Mueller§, and Nancy B. Schwartz†§

From the Departments of †Pediatrics, §Pathology, and ‡Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637

Expression of the extracellular proteoglycan aggrecan is both cell-specific and developmentally regulated. Previous studies identified six functionally defined cis elements in the aggrecan promoter region which were shown to repress aggrecan gene expression (1). Using competition electrophoretic mobility shift assays (EMSAs) we have now identified in nuclear extracts a functional repressor cis element, (T/C)TCCCCT(A/C)RRC, which occurs at multiple locations within the chick aggrecan regulatory region. We purified the factor that binds to this cis element and established that it, APBP-1 (aggrecan promoter-binding protein-1), is a 19-kDa protein that has significant homology to CIRP (cold inducible RNA-binding protein). Recombinantly expressed APBP-1 mimics the native cis element-trans factor interaction in EMSAs. In situ hybridization demonstrates that aggrecan and APBP-1 RNA expression are restricted to complementary tissues in the developing limb, and Northern blot analysis of chick limb bud mRNA shows that APBP-1 mRNA expression is inversely correlated with aggrecan mRNA expression. Functional analyses by transient transfections and Northern blot analyses suggest APBP-1 has the capacity to repress aggrecan expression, indicating that this factor may be important regulator of aggrecan gene expression.

Differentiation of mesenchymal cells to chondrocytes and elaboration of those cells of an extracellular matrix composed of aggrecan, link protein, and collagen types II, IX, and XI results in the highly specialized connective tissue cartilage, which exhibits unique biochemical and biomechanical properties. Coordinate regulation of the genes underlying these properties may be essential for normal skeletal development and subsequent maintenance of the cartilage phenotype in postnatal life (2).

Eukaryotic gene regulation entails interplay between activators and repressors of transcription whose interactions influence the tissue- and developmental stage-specific expression of multitudes of genes. The aggrecan gene is dynamically regulated throughout embryonic life, with both mRNA and its protein product being up-regulated by 50-fold from day 2 to day 6, then undergoing a 2-fold decline by day 8 and reaching a steady plateau thereafter (2). Disruption of the delicate balance of transcription factors during development and/or later in life may lead to pathologic conditions. Although the etiologies of osteoarthritis and rheumatoid arthritis remain unknown, aggrecan degradation as well as reduced de novo matrix production is associated with each of these disease states.

Therefore, to understand better how aggrecan gene expression is regulated, we have characterized the chick aggrecan gene promoter and enhancer region with the aim of identifying regulatory elements and their cognate transcription factors. We mapped a 400-bp region that actively mediates repression of aggrecan expression, removal of which increases aggrecan promoter activity by nearly 3-fold (3). Three DNase I-protected footprint sequences that mediate repressor activity in the context of this 400-bp sequence have been identified and characterized (1). The three protected sequences, named J (nucleotides -878 to -831), K (nucleotides -808 to -791), and L (nucleotides -730 to -703), contain various combinations of the sequences CACCTCC, TTCCTCC, and CTCTCC. Mutation of each of these sites singly yielded nearly a 3-fold increase in promoter and enhancer activity, and each of these protected sequences produced a unique electrophoretic mobility shift assay (EMSA) pattern that is cell type-specific (1). Competition EMSAs revealed that multiple tissue-specific DNA-protein complexes exist and that one complex in particular, designated band C, is detected with probes that contain these sequences. Although no transcription factor(s) acting at these sites has been described, the sequence elements CACCTCC, TTCCTCC, and CTCTCC have also been shown to be important to varying degrees in the regulation of other matrix-specific genes (4–6).

In addition to the data indicating that the above mentioned sequences contribute to the regulation of the chick aggrecan gene (1), a subset of these sequences has been shown to be associated with the transcription factors Sp1, Sp3, and ßEF1 in the context of the human COLIA2 and COL2A1 genes and the rat Col2a1 gene (4–7). In all of these studies, EMSA analysis revealed the presence of multiple complexes; therefore, there probably are other yet-to-be identified transcription factors that contribute to the regulation of matrix genes via these sequences. The presence of these sequence elements in the chick aggrecan promoter and enhancer region (1) as well as in other matrix gene regulatory regions (4–6) raises the possibility that each sequence could contribute singly and/or in concert to cartilage-specific regulation of expression. Thus, our goal in this study was to characterize further one recently identified cis element that is functional in the chick aggrecan promoter and to use its sequence as a target to identify the interacting protein(s) present in chick cartilage. Consequently, a unique RNA/DNA-binding protein has been isolated, characterized, and shown to regulate aggrecan expression.

*This work was supported by United States Public Health Service Grants HD-17332 and HD-09402 (to N. B. S.) and Training Grants HL-07237 and HD-007009 (to E. W. P.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence reported in this paper has been submitted to the GenBank™/Ebi Data Bank with the accession number(s) DQ126684.

1 To whom correspondence should be addressed: Dept. of Pediatrics, The University of Chicago, 5841 S. Maryland Ave., MC 5058, Chicago, IL 60637. Tel.: 773-702-6426; Fax: 773-702-9234; E-mail: n-schwartz@uchicago.edu.

2 The abbreviations used are: EMSA, electrophoretic mobility shift assay; APBP-1, aggrecan promoter-binding protein-1; CIRP, cold inducible RNA-binding protein; DIG, digoxigenin; EST, expressed sequence tag; GST, glutathione S-transferase; HPLC, high performance liquid chromatography.
EXPERIMENTAL PROCEDURES

**Materials**—Oligonucleotides were made with an Applied Biosystems 3808 DNA Synthesizer. Reagents for biochemical and molecular cloning experiments were of the highest quality available from commercial vendors. Spl and Sp3 antibodies were purchased from Santa Cruz Biotechnology, Inc., Santa Cruz, CA.

**Cell Culture**—Cultures of day 14 chick sternal chondrocytes were established as described previously (1, 3); however, some cells grown for protein purification were incubated in 75-cm² flasks rather than in dishes. Chondrocyte cultures were established in F-12 medium and fed on E14C0 and E14C4 with either F-12 or Dulbecco’s modified Eagle’s medium supplemented with 10% fetal calf serum, as described under “Results.”

**Transient Transfection**—Conditions of transfections conducted for luciferase and Northern blot experiments were as described by Pirok et al. (1, 3) using the calcium phosphate method. Triplicate plates containing ~5 × 10⁶ cells (day 12 chondrocytes)/100-mm tissue culture dishes (Falcon) received between 10 and 20 pmol of a given plasmid construct to be assayed. One μg of a β-galactosidase reporter plasmid was cotransfected with each experimental construct to correct for cell loss. The transfections were allowed to proceed for 36 h.

For *in situ* transfections, chondrocyte cultures established from E12 sternal cartilage were treated with 1 μg of purified recombinant APBP-1 using the Chariot™ Advantage delivery reagent as a vehicle (+ APBP-1) or with 1 μg of albumin (−APBP-1). 18 h after treatment the cells were fixed, and *in situ* hybridization was performed using digoxigenin (DIG)-labeled aggrecan probe.

**EMSA**—100 pmol of each single-stranded DNA oligonucleotide was incubated in annealing buffer (100 mM Tris, pH 7.5, 500 mM NaCl, and 250 mM MgCl₂) at 90°C for 5 min and allowed to cool slowly for approximately 1 h. Standard methods were employed to end label and purify the double-stranded oligonucleotides. Briefly, 10 pmol of probe was incubated for 30 min at 37°C with kinase buffer (Promega), 20 units of T4 kinase, and [γ⁻³²P]ATP (PerkinElmer Life Sciences). To stop the labeling reaction 0.5M EDTA was added to the reaction, and the labeled double-stranded oligonucleotides were purified by HPLC experiments. 200 pmol of probe was added to the reaction and annealed in renaturation buffer (100 mM Tris, pH 7.5, 50 mM NaCl, 2 mM EDTA, 1 mM dithiothreitol, 0.2% Tween 20 [w/v], 30 mM KCl), then DNA-binding probes were eluted stepwise with increasing concentrations of KCl: 500 mM, 1 M, 2.5 M, 3.2 M, and 4 M. The wash and elution fractions were dialyzed against 10 mM ammonium bicarbonate, pH 7, and lyophilized, then resuspended with 10 mM Tris, pH 8, the proteins separated by 15% polyacrylamide SDS-PAGE, and the gel was stained with Coomassie Blue. Typically 5–20 μg of protein was recovered/experiment.

**Mass Spectral Analysis and Microsequencing**—Mass spectral analysis, trypsin digestion, ProFound and MS-FIT data base searches, and microsequencing of protein specimens were performed by the Rockefeller University Protein/DNA Technology Center. ProFound is accessible at prowl.rockefeller.edu/cgi-bin/ProFound; MS-Fit is available from the University of California San Francisco at prospector.ucsf.edu/

**DNA Cloning and Protein Purification**—The Invitrogen ProEXHTa system, which introduces His₉ residues at the N terminus of the expressed protein, was used for bacterial expression of the cloned putative chick APBP-1 with the following primers: A1(Ncol), 5'-CATGCACTGGCCTGATAGGCC-3'; and A1.Xbal), 5'-GCATCTAGATTACCTGTGTGTTAGCCATACCTG-3'. For subcloning into the ProEXHTa vector, PCR-generated DNA fragments were digested with Ncol and Xbal, then ligated into the vector that had been digested with the same restriction enzymes. All clones were sequenced in their entirety before transformation into JM109 DE3 cells (1, 3). For mammalian expression experiments the pTarget (Promega) vector was used according to the manufacturer’s protocol. Briefly, the ProEXHTa vector, or the APBP-1 insert was used as a template for a PCR in which the primers Cirp1 5'-ATGGCATCGGATGAGGGCAAG-3' and A1.1(XbaI), 5'-GACATCTAGATTACCTGTGTGTTAGCCATACCTG-3' were used to generate a PCR fragment that was subcloned into the pTarget vector. The clones were sequenced to ensure that no errors were generated during the amplification or cloning procedures. Expressed protein from bacteria was purified as described (8). Briefly, overnight bacterial cultures were centrifuged and the resultant pellet sonicated in buffer containing 5 mM imidazole and 50 mM Tris, pH 7.9. After removal of bacterial debris by ultracentrifugation, the supernatant was loaded onto a 5-ml affinity column of nitrilotriacetic acid resin charged with nickel (Qiagen), which binds to the imidazole moieties of the His tag residues to capture the expressed protein. Columns were
TABLE ONE

| Name  | Protected sequence                                      | Migratory complex name | Positions |
|-------|--------------------------------------------------------|------------------------|-----------|
| H     | 5‘-CTGTTAGGGAGACTCTCTGAGA-3’                         | A, C                   | −1015 to −996 |
| I     | 5‘-CTGCCCACACTGACCTTCCAGG-3’                         | 1, 2                   | −960 to −937 |
| 1.2   | 5‘-TCCCCACCTCTCTCCAG-3’                                | 1, 2                   | −878 to −861 |
| 1.1   | 5‘-CTTCCCACAGCCCTCCAGAGGACCCAGAG-3’                   | 1, 2, 3, 4             | −860 to −831 |
| K     | 5‘-CTGGGCTTCCCTCCAGAG-3’                               | A, C                   | −808 to −791 |
| L     | 5‘-CTGTTGACACCCCTCCCTCAGAG-3’                         | 1                      | −730 to −703 |

*Names of the sequences used for EMSA probes were derived from the protected sequences determined via previous DNase I footprinting experiments of the aggrecan promoter and enhancer region. Relative positions are in reference to the first 5’-transcriptional start site of the chick aggrecan promoter.*

RESULTS

We previously reported the identification and characterization of three DNase I-protected footprint sequences, designated J, K, and L, which span the positions −878 to −831, −808 to −791, and −730 to −703, respectively, and which mediate repressor activity in the context of the chick aggrecan promoter and enhancer region (TABLE ONE) (1). These three protected regions contain several similar and repetitive sequences, including repeats of the nucleotide sequences TCCCC and CTCTCCC. Comparison of EMSAs using crude nuclear protein from chick sternal chondrocytes, hybridized with probes J.1, J.2, and L, which correspond to DNase I-protected sequences J and L, revealed one complex, band C, which was produced by three different probes with similar sequences containing the subsequence TCCCC, and another complex, band B, that was produced by two of these three probes (TABLE ONE) (1). Another sequence repeated within these probes was TCCCTCC, a sequence that has been shown to bind to Sp1 or Sp3 (5). The CTCTCCC sequence was deduced to be a component of complex A (TABLE ONE) (1).

Because crude nuclear protein bound EMSA probe L with greater affinity and specificity than some of the other probes containing similar sequences, this sequence was used to examine the nature of the functional *cis*-element-*trans* factor interaction observed previously. We also examined whether Sp1 or Sp3 was participating in binding to any of these sequences or complexes using crude nuclear protein derived from day 14 chick sternal chondrocytes. The specific goals of the present studies were 6-fold: 1) to define which protected nucleotides of sequence L are necessary and sufficient to bind repressors of aggrecan expression in primary chondrocytes; 2) to determine whether Sp1 or Sp3 is involved in binding to the protected sequences K, J, or L; 3) to determine what other *trans* factor(s) is binding to this sequence; 4) to clone the transcription factor that binds to the core L sequence and to characterize its DNA binding properties; 5) to examine the expression patterns of the identified transcription factor in relation to aggrecan message levels *in vivo* and *in vitro* developmental model; and 6) to determine whether recombinant expression of the *trans* factor can functionally repress aggrecan expression.

Core Sequence Necessary for Nuclear Protein Binding in the Protected Region L—To identify the nucleotides in region L necessary for the observed repressor activities, competition EMSAs with extracts derived from primary chondrocytes and probes bearing triple-substitution mutations of the native sequence (TABLE TWO) were used to define the minimal sequence necessary for the region L DNA–protein interactions (Fig. 1). Competition with 20 fmol of radiolabeled probe L showed that complexes B and C can be competed out with the mutated unlabeled oligonucleotides (400 fmol) L1, L3, L4, and L9 but not by L2, L5, L6, L7, or L8 (Fig. 1). The results for L5, L6, L7, and L8 suggest that motifs in the L subsequence TCCCCCTCAGCA are necessary for formation of complexes B and C. The inability of mutated L2 to compete for bands B and C suggests that mutation of the CTT triplet changed a half-site required for proper binding of the transcription factors to the primary site located in the TCCCCCTCAGCA sequence.

Taken together with our earlier observation that probe J.1 (TABLE ONE) contains the sequence TCCCCCTAGAC and also forms com-
plexes B and C (1), a deduced consensus sequence for complex B and C binding is (Y)TCCCCT(A/C)RRC, a novel cis element. Interestingly, another probe, J.2 (TABLE ONE), forms complex C but not B, and has a shortened version of this consensus sequence, TTCCCC (1). Perhaps some or all of the nucleotides T(A/C)RRC in the 3′-portion of the consensus sequence are required for the formation of complex B, or an upstream sequence containing CTT is necessary within 5–6 bases of the (Y)TCCCCT(A/C)RRC core for proper formation of complex B.

Potential Sp1/Sp3 Binding to Chick Aggrecan Promoter—Because it was reported previously that Sp1 and Sp3 have the ability to bind to a TCCTCC consensus sequence present in the COL1A2 promoter (5), we examined whether any interactions occur between anti-Sp1 or anti-Sp3 antibodies and EMSA complexes formed with probes K and J.2, which both contain the TCCTCC sequence. Fig. 2 shows the results of incubating crude nuclear protein derived from chick sternal chondrocytes with probes J.1, J.2, K, and L (defined in TABLE ONE). Fig. 2A demonstrates that the addition of anti-Sp1, anti-Sp3, or a combination of anti-Sp1 and anti-Sp3 antibodies had no effect on migration of complexes formed by probe J.2, which contains the TCCTCC motif. Addition of anti-Sp3 or the combination of anti-Sp1/Sp3 to probe J.1 complexes did cause a slight decrease in band D and slight increases in bands B and C. As controls, unlabeled probe J.1 competed for bands B, C, and D formation, and cold unlabeled probe J.2 competed for formation of bands A and C. Fig. 2B presents the results of assays with probes K and L. Probe K, which contains a TCCTCC sequence, generated a doublet with only the upper band competed for by excess unlabeled probe K. Addition of anti-Sp1, anti-Sp3, or a combination of anti-Sp1 and anti-Sp3 at this, or other lower concentrations of antibody (data not shown), had no visible effect on DNA-protein complex size. As positive controls, supershift assays with Sp1/Sp3 antibodies were done utilizing HeLa extracts (Promega) and 32P-radiolabeled double-stranded DNA probes that corresponded to the published consensus sequence of Sp-1 5′-ATTCCGATCGGGGCGGGCGAG-3′ (Promega) and Sp-3 5′-CTCTGC-
protein(s) binding to the L footprint sequence. A cluster of L-binding fractions was observed in the Southwestern analysis of the HPLC fractions was performed using the L sequence as probe. A cluster of L-binding fractions was observed in the 19–21-kDa range; the most prominent were present in fractions 43–45. Southwestern analysis of these fractions revealed a prominent band corresponding to a mass of ~19 kDa from these fractions. As a control, crude nuclear protein not purified via HPLC resulted in a band at approximately the same size and a less intense larger species, indicating that sequence L may bind a multiprotein complex in vivo.

To obtain sufficient amounts of the protein(s) binding to the L sequence to permit microsequencing and identification, a second method of purification was employed: Centricon fractionation followed by magnetic bead affinity chromatography. After loading of the sample, the magnetic affinity beads were washed three times with binding buffer, designated as fractions W1–W3; the column was then eluted stepwise with increasing concentrations of KCl: 1, 2.5, 3.2, and 4 M, designated as E1, E2, E3, and E4, respectively. Fig. 4A shows the result of this purification scheme, revealing a doublet of protein bands at about 19–21 kDa in the elutions. When a duplicate isolation was conducted and the gel was subjected to an in-gel Western procedure using radiolabeled probe L in the presence of sonicated salmon sperm genomic DNA (Fig. 4B), the smaller band of the doublet was labeled, showing that this 19-kDa protein is capable of binding to probe L.

To ensure that the purified protein was not an artifact of tissue culture, the above purification scheme was repeated with nuclear extracts from E14 chick sternal cartilage (Fig. 4C). Under all conditions, the resultant Coomassie-stained gels exhibited a doublet at 19–21 kDa, indicating that it is native to the fully developed and differentiated tissue and is present in quantities sufficient to purify. Because probe L has the tendency to form two distinct complexes (B and C) with crude nuclear proteins (Fig. 1), EMSAs were performed with the purified nuclear protein to identify which of the two bands (B or C) corresponded to the
19-kDa protein. As seen in Fig. 4D, lanes W1–W3, the washes did not exhibit a discernible band, whereas the elution at 1 M KCl showed a gel shift band that lines up with complex C (compare lane E1 with lane C). This suggests that other proteins may interact with probe L to form band B or that other proteins from the nuclear extract interact with complex C to form B. Taken with the results that the Coomassie-stained gel exhibits a protein doublet at ∼19–21 kDa and the same protein is reactive with probe L by Southwestern analysis, this indicates that complex C is produced by the 19–21-kDa protein purified from day 14 chick sternal chondrocytes, which we now designate as APBP-1.

Both bands of the 19–21-kDa doublet were excised separately from 15% SDS-polyacrylamide gels and subjected to trypsin digestion, mass spectrometric analysis, and microsequencing. Results from mass spectrometric analysis (Fig. 5A) were used in ProFound and MS-FIT data base searches, tentatively identifying both proteins of the 19–21-kDa doublet as similar to mouse CIRP (10). Microsequencing of cleaved peptides from the smaller protein of the doublet identified the 19-kDa protein as a chick version of the family of CIRPs (Fig. 5B). Currently, there have been no functional studies on this family of proteins to investigate potential involvement in DNA binding or gene regulation. The fact that the slightly larger band copurifies with the functional product raises the possibility that the second band could be a differently processed form of the 19-kDa protein.

Cloning Chick APBP-1—To obtain the full chick sequence, BLAST searches of the Roslin Institute chicken expressed sequence tag (EST) and NCBI nonmouse EST data bases with the mouse CIRP sequence were performed. One of these, A1394576 in the NCBI data base EST, encodes a full-length counterpart of the mammalian CIRPs, including a match of a 10-amino acid residue C-terminal sequence obtained from our protein isolate. Oligonucleotide primers derived from this EST sequence successfully amplified a cDNA by reverse transcription-PCR using chick sternal chondrocyte total RNA as template. Sequencing showed that this cDNA contains a 516-bp open reading frame encoding a 172-amino acid residue sequence; comparison of the chick cloned and EST DNA-predicted peptide sequences found 99% identities. Similarities of APBP-1 and other CIRPs are listed in Table 3. Furthermore, with this sequence we have confirmed the mass of at least five of the six peaks detected in our original mass spectra matches. The sixth peak could be matched using the FindMod tool (us.expasy.org/tools/findmod/) to a phosphorylated tryptic peptide that contains at least two serine residues predicted to be phosphorylated by the NetPhos 2.0 Server (www.cbs.dtu.dk/services/NetPhos/). Moreover, searches of chick EST data bases (www.genome.ucsc.edu/cgi-bin/hgBlat) also revealed an alternative splice form of the C-terminal region differing from our APBP-1 sequence. The open reading frame predicts a protein of 20973.94 Da and provides an explanation for the 21-kDa protein obtained in our original purification.

The chick APBP-1 cDNA was inserted into the vector ProEx HT (Invitrogen), expressed in Escherichia coli, and the protein recovered by His tag affinity chromatography. The predicted peptide sequence yielded a calculated molecular mass of about 18.6 kDa, whereas the recombinant His tag protein had an apparent size of about 20 kDa by SDS-PAGE (data not shown); both estimates are similar to the 19-kDa size determined for the isolated L-binding protein. EMSAs with crude nuclear protein from day 14 chick chondrocytes and recombinant chick APBP-1 were performed to determine DNA binding properties. EMSAs were conducted with probes H, I, J1, J2, K, and L and the cloned chick APBP-1 protein, with the addition of histidine residues at the N terminus, referred to as (His)6APBP-1. Complexes of DNA and recombinant APBP-1 formed after incubation with probes J1 and L but not with probes H, I, J2, or K (Fig. 6); both probes J1 and L contain the sequence TCCCT(A/C)RRC. By all characteristics examined to date, the cloned protein appears to have DNA binding properties identical to APBP-1, the L-binding protein isolated from chick nuclear protein.

Expression of Chick APBP-1 during Development—Our laboratory previously demonstrated modulation of aggrecan expression in early development of the chick, but the mechanism of regulation is unknown. Chick chondrocytes contain aggrecan mRNA at embryonic day 12 and at late embryonic day 14, but not at embryonic day 14. Of particular interest, a band at ∼20 kDa is present at embryonic day 14 that is not present at embryonic day 12 (11). This band is the same size as the band shown in Fig. 4C, lane 9, of the chick chondrocyte nuclear protein. For those experiments, we isolated nuclear proteins from embryonic day 14 chick sternal chondrocytes without culturing. The crude nuclear protein from day 14 chick sternal chondrocytes, which we now designate as APBP-1, the L-binding protein isolated from chick nuclear protein.
limb bud development from very low levels before embryonic day 4 to a ~50-fold greater mRNA and protein levels by day 5, followed by down-regulation of aggrecan message and protein levels as mesenchymal condensation is established (2). To assess quantitatively the relationship between APBP-1 and aggrecan expression, we employed limb bud cultures derived from day 2 embryos, which faithfully mimic the in vivo aggrecan regulation pattern. During the developmental progression in this model system, we analyzed APBP-1 and aggrecan expression by Northern blots. Up to day 4 there is negligible aggrecan mRNA expression by these cultures (Fig. 7) and no deposition of aggrecan in the extracellular matrix. By day 5 there is a marked increase in aggrecan mRNA expression, followed by a decrease in expression by day 7 (Fig. 7), as documented previously on the transcript and protein levels using aggrecan probes and the monoclonal antibody S103L, respectively (2). In contrast, APBP-1 expression follows an inverse pattern of expression compared with aggrecan, with initial high levels of APBP-1 expression and a substantial decline by day 5, concomitant with aggrecan message levels being established. By day 7, there was a slight but significant

FIGURE 5. Mass spectroscopy of L-affinity-purified protein and amino acid sequence from the 19-kDa protein. A is the mass spectral analysis output of tryptic peptides resulting from the L-affinity-purified 19-kDa protein from chick chondrocytes isolated by gel electrophoresis and band excision. B, microsequencing of one of the tryptic fragments of the 19-kDa protein yielded the indicated sequence. For comparison, a schematic map of mouse CIRP is provided, and the location of the sequence similarity is shown.
To determine the distribution of chick APBP-1 in the developing limb, mRNA in situ hybridizations were performed on limb sections from day 12 chick embryos. Aggrecan mRNA expression is restricted to developing cartilage; in fact, the localization pattern shows that aggrecan mRNA is more highly expressed in prehypertrophic chondrocytes in the growth plate, and as differentiation proceeds aggrecan mRNA levels are strongly down-regulated in hypertrophic chondrocytes (Fig. 8A). In contrast, APBP-1 expression is strongest in the surrounding tissues such as the adjacent muscle (M), where aggrecan expression is absent (Fig. 8B). Within the growth plate, early prehypertrophic chon-
APBP-1, a Transcriptional Regulator of Aggrecan

FIGURE 9. Aggrecan expression and promoter activity after treatment with exogenous chick APBP-1. A shows in situ analyses of chick chondrocytes after APBP-1 or control transfections. Chondrocyte cultures established from E12 sternal cartilage were treated with 1 μg of purified recombinant APBP-1 using the Chariot™ Advantage delivery reagent as a vehicle (+ APBP-1) or with 1 μg of albumin (− APBP-1). 18 h after treatment the cells were fixed, and in situ hybridization was performed using DIG-labeled aggrecan mRNA probe. Representative pictures are shown. B shows a Northern blot comparing the aggrecan message levels of day 14 chondrocytes cultured with or without 10 μg of transfected APBP-1. 10 μg of total RNA was extracted from transfected day 14 chick sternal chondrocytes. RNAs in lanes 1 and 2 were purified from cultures that were either transfected with 10 μg of empty pTARGET vector or with 10 μg of APBP-1-driven pTARGET, respectively. Blots were probed with a 1.25-kb probe that corresponding to the chick aggrecan chondroitin sulfate exon, which was 32P-labeled using the Rediprime II kit. The blot was then stripped and reprobed with a probe of the chick housekeeping gene for 18 S RNA. C and D are cotransfections of APBP-1-expression and aggrecan promoter-reporter plasmid constructs. Luciferase reporter activities were assayed after cotransfection of cultured chick chondrocytes with the indicated plasmid construct combinations. In both C and D Ag-1(+) is the 1.8-kb promoter/enhancer region from the aggrecan gene placed in the reporter vector pGL2-Basic, whereas Basic is the reporter vector pGL2-Basic with no insert. pTARGET was the empty expression vector pTARGET with no cDNA inserted. APBP-1 was the pTARGET vector with chick APBP-1 cDNA inserted in the expression site.

chondrocytes, in which expression of aggrecan is maximal (Fig. 8B), do not express APBP-1 (Fig. 8E); whereas in the interface between the prehypertrophic and hypertrophic regions there is a band of APBP-1 expression that could contribute to down-regulating the high levels of aggrecan expression exhibited by the prehypertrophic chondrocytes (Fig. 8D). In the mature hypertrophic chondrocytes (Fig. 8C) small but detectable levels of APBP-1 are observed. The expression patterns are consistent with APBP-1 being involved in aggrecan regulation during growth plate differentiation in vivo. As observed by Northern blot analysis, APBP-1 expression exhibits an inverse relationship with that of aggrecan both temporally and spatially. These are the first studies that examine expression of chick APBP-1 in vivo, and they provide a foundation upon which to examine the mechanisms by which APBP-1 may regulate aggrecan expression.

Effect of Exogenous Chick APBP-1 on Aggrecan Expression—To determine whether recombinant chick APBP-1 can attenuate aggrecan message levels, purified (His6)APBP-1 protein was delivered to chondrocyte cultures that were established from day 12 embryos using Chariot™ Advantage as a carrier. Cells were incubated for 24 h and then probed for aggrecan message by mRNA in situ hybridization. Control chondrocytes, to which albumin was delivered, exhibited relatively high expression of aggrecan in the individual proliferating chondrocytes (Fig. 9A). In contrast, APBP-1-treated chondrocytes exhibit a decrease in aggrecan expression, suggesting that rapid down-regulation of aggrecan expression follows the addition of recombinant APBP-1. Furthermore, chondrocytes transfected with an APBP-1-expressing construct exhibit quantitatively reduced levels of aggrecan message, as determined by Northern blot analysis (Fig. 9B), suggesting that APBP-1 has the capacity to reduce aggrecan message levels specifically. Together, these results suggest that APBP-1 is functionally active in repressing aggrecan expression in culture.

To determine whether recombinant APBP-1 affects the transcriptional activity of the aggrecan promoter, a luciferase reporter construct containing the 1.8-kb Ag-1(+) (3) chick aggrecan promoter sequence was cotransfected with an APBP-1 expression construct into cultured chick sternal chondrocytes. Significant reduction of luciferase activity was observed when APBP-1 was cotransfected with the Ag-1(+) construct (Fig. 9C). Cotransfections of increasing amounts of the APBP-1 construct with the Ag-1(+) construct revealed that repression of luciferase reporter activity was dependent on the amount of exogenous APBP-1 expression plasmid added to the culture (Fig. 9D). In contrast, the absence of APBP-1 expression vector had no significant impact on background luciferase expression from the empty reporter vector (Fig. 9C), implying that the effect of APBP-1 on Ag-1(+) reporter expression is at the transcriptional level.
DISCUSSION

Overall, this study has established the minimal core sequence of a functional repressor sequence located within the chick aggrecan promoter and enhancer region, a sequence that has been proposed to be important in the context of other cartilage-specific genes. We have also identified a putative transcription factor binding to this specific sequence, thereby elucidating both components of a cis-element-trans factor interaction bearing on repression of the chick aggrecan promoter. The sequence (T/C)TCCCCCT(A/C)RRC has the capacity to form two major protein-DNA complexes, but it does not interact with Sp1 and Sp3 in nuclear extracts from primary cultures of chick chondrocytes. It does, however, bind to a novel 19-kDa transcription factor named APBP-1, which has homology to mouse CIRP. We have cloned the chick homolog of CIRP, APBP-1, and have determined that it is expressed in a pattern complementary to that of aggrecan and that it has the capacity to repress transcription initiated from the aggrecan promoter and enhancer region.

The isolated chick APBP-1 cDNA encodes a protein of the same length as mouse CIRP with 94% amino acid sequence identity. Mouse CIRP is an 18-kDa protein composed of 172 amino acids (10) and is classified as an RNA-binding protein because of the presence of a ribonucleoprotein motif at the N terminus. Increasingly, members of this family of highly similar RNA-binding proteins have been identified from many species, ranging from Xenopus to humans (10–14). The Xenopus homolog of CIRP, named xCIRP2, and mammalian CIRP were shown to localize to the nucleus in cultured cells and xCIRP2 exhibited nucleocytoplasmic shuttling (15). Proteins that exhibit the ribonucleoprotein motif have been shown to regulate gene expression post-translationally by binding to an mRNA and affecting its stability, translation, or splicing (16). Northwestern analysis of a GST-CIRP fusion protein showed that the protein could bind homopolymers of the nucleotides A, C, G, and U, and as expected, the greatest affinity was for poly(U) (10). Overexpression of this protein, which can be induced by exposure to low temperature, leads to G1 prolongation and reduced growth rate (10). The results reported here for probe L and the putative chick APBP-1 are the first to show a CIRP-like protein binding to a specific DNA sequence in any organism.

This novel DNA-binding protein may also be involved in the regulation of other genes in which similar cis elements have been reported. The site CACCTCC has been linked with a specific transcription factor in the context of the rat Col2a1 gene (4). EMSAs on the promoter revealed formation of a minimum of four DNA-protein complexes. One of the four complexes was identified as δEF1 (4). The remaining three complex components as yet have not been identified; one of the unidentified three could be APBP-1 because it binds to the rat Col2a1 identical sequence upstream of the δEF1 binding site. In that study, a probe containing the silencer sequence found at bp 439 of the rat type II collagen promoter also bound δEF1, suggesting that δEF1 also has the capacity to bind to a native sequence in the rat Col2a1 promoter. We found that this probe contained a sequence motif also present in the chick aggrecan promoter and enhancer region. The actual probe sequence from the rat Col2a1 promoter was AACTCCCCATC-CCCACCTCCTTTCCTCCC, with the underlined nucleotides representing the δEF1 binding site (4), the bold nucleotides representing the motif we have found in the chick aggrecan promoter footprints J and L (TCCCC), and the italics representing another motif present in our probes K and J (CTCCC). Because these matrix-specific genes all contain conserved cis elements in their regulatory region, there may be a common trans factor complex that interacts with these individual regulatory regions.

APBP-1, a Transcriptional Regulator of Aggrecan

Expression of recombinant chick APBP-1 revealed that it has the capacity to bind to core DNA sequences in a fashion identical to the crude nuclear protein constituent with specificity for this sequence. In situ experiments revealed that APBP-1 is expressed in tissues where aggrecan is not being actively synthesized, a pattern consistent with its presumed function as a transcriptional repressor of the aggrecan promoter in vivo. Northern blot analysis in a developmental limb bud culture system demonstrated that APBP-1 is expressed in a temporal manner that is inversely correlated to aggrecan message levels. Experiments that address the question of a direct functional relationship between aggrecan and APBP-1 show that introduction of APBP-1 protein or message to cells effectively reduces matrix production or aggrecan message levels. Furthermore, in a luciferase-reporter aggrecan-promoter assay system, transfection with APBP-1 reduces reporter activity in a dose-dependent manner. The fact that this protein was originally classified by its RNA binding properties makes it tempting to speculate that this transcription factor is regulating the aggrecan promoter, and perhaps other genes, by binding to both DNA, as we have shown, and perhaps to RNA during the process of translation. Preliminary experiments in our laboratory have shown that purified APBP-1 has mRNA binding capacities, and competition binding experiments with DNA do not disrupt the RNA-protein interactions (data not shown). APBP-1 and CIRP may be members of a subfamily of proteins that function to regulate both the transcription and translation of critically important, developmentally regulated genes.

The dual nature of RNA-binding proteins regulating transcription is an emerging field. Examples of proteins that can provide dual functionality (17) include TFIIIA, which can bind and regulate both DNA and RNA (18); binding to RNA is dependent on secondary and tertiary structure rather than primary sequence, whereas binding to DNA is specific to the primary structure of DNA (19). The homeodomain protein bicoid is another example of a protein that has the capacity to regulate both DNA and RNA (20); helix III, the homeodomain, is responsible for both DNA and RNA regulation but does so via separate functional protein domains. Regulation of alternative splicing by ribonucleoprotein domain-containing proteins has also been suggested in the literature (21). Aggrecan has been shown to be alternatively spliced for the epidermal growth factor (EGF)-like domains, but the functional consequences of this are not clear except for differences in the ability to interact with other extracellular matrix molecules (22, 23). Thus, it is possible that APBP-1 might function in the splicing/processing of aggrecan mRNA as well. Future work will focus on investigating whether DNA and RNA binding occur at separate or shared functional protein domains, as it is becoming more evident that gene regulation is influenced not only by classic cis elements-trans factors interactions but also by RNA processing both within and outside of the nucleus.

It is intriguing that the 19-kDa protein that binds to probe L and produces complex C is related to this relatively novel family of proteins, especially if it is found to be modulated by stress, as it has been reported that CIRP is down-regulated by hypoxia (24) in other systems/tissues. One would predict that the activity of the aggrecan promoter should be sensitive and responsive to external stress, logically via a transcriptional regulator modulated by stress, whether it be mechanical, such as the load on cartilage in weight-bearing joints, or hypoxia, a physiological stress to which anaerobic chondrocytes must adapt. Hence, the identification of APBP-1 will enable further investigation of the mechanisms in which aggrecan, and perhaps other matrix-specific genes, are regulated during development and in disease states in which there is limited matrix production.
Acknowledgment—We thank James Mensch for advice and manuscript review.

REFERENCES

1. Pirok, E. W., III, Henry, J., and Schwartz, N. B. (2001) J. Biol. Chem. 276, 16894–16903
2. Schwartz, N. B., Hennig, A. K., Krueger, R. C., Krzystolik, M., Li, H., and Mangoura, D. (1993) in Limb Development and Regeneration (Fallon, J. F., Goetinck, P. F., Kelley, R. O., and Stocum, D. L., eds) pp. 505–514, Wiley-Liss, Inc., New York
3. Pirok, E. W., III, Li, H., Mensch, J. R., Henry, J., and Schwartz, N. B. (1997) J. Biol. Chem. 272, 11566–11574
4. Murray, D., Precht, P., Balakir, R., and Horton, W. E., Jr. (2000) J. Biol. Chem. 275, 3610–3618
5. Ihn, H., LeRoy, E. C., and Trojanowska, M. (1997) J. Biol. Chem. 272, 24666–24672
6. Ihn, H., Ohnishi, K., Tarnaki, T., LeRoy, E. C., and Trojanowska, M. (1996) J. Biol. Chem. 271, 26717–26723
7. Ghayor, C., Chadjichristos, C., Herrouin, J. F., Ala-Kokko, L., Suske, G., Pujol, J. P., and Galera, P. (2001) J. Biol. Chem. 276, 36881–36895
8. Deypur, A. T., Krishnan, S., Cockburn, B. N., and Schwartz, N. B. (1998) J. Biol. Chem. 273, 9450–9456
9. Grove, E. A., Tole, S., Limon, J., Yip, L., and Ragsdale, C. W. (1998) Development 125, 2315–2325
10. Nishiyama, H., Itoh, K., Kaneko, Y., Kishishita, M., Yoshida, O., and Fujita, J. (1997) J. Cell Biol. 137, 899–908
11. Nishiyama, H., Higashitsuji, H., Yokoi, H., Itoh, K., Danno, S., Matsuda, T., and Fujita, J. (1997) Gene (Amst.) 204, 115–120
12. Maruyama, K., Sato, N., and Ohta, N. (1999) Nucleic Acids Res. 27, 2029–2036
13. Saito, T., Sugimoto, K., Adachi, Y., Wu, Q., and Morii, K. J. (2000) Comp. Biochem. Physiol. B Biochem. Mol. Biol. 125, 237–245
14. Tanaka, K. J., Kawamura, H., and Nishikata, T. (2000) Dev. Genes Evol. 210, 464–466
15. Aoki, K., Ishii, Y., Matsumoto, K., and Tsujimoto, M. (2002) Nucleic Acids Res. 30, 5182–5192
16. Burd, C. G., and Dreyfuss, G. (1994) Science 265, 615–621
17. Mayfield, S. (1996) Chem. Biol. 3, 415–418
18. Clemens, K. R., Wolf, V., McBryant, S. J., Zhang, P., Liao, X., Wright, P. E., and Gottesfeld, J. M. (1993) Science 260, 530–533
19. McBryant, S. J., Veldhoen, N., Gedulin, B., Leresche, A., Foster, M. P., Wright, P. E., Romaniuk, P. J., and Gottesfeld, J. M. (1995) J. Mol. Biol. 248, 44–57
20. Dubnau, J., and Struhl, G. (1996) Nature 379, 694–699
21. Auboeuf, D., Dowhan, D. H., Li, X., Larkin, K., Ko, L., Berget, S. M., and O’Malley, B. W. (2004) Mol. Cell. Biol. 24, 442–453
22. Doege, K. J., Sasai, M., Kimura, T., and Yamada, Y. (1991) J. Biol. Chem. 266, 894–902
23. Day, J. M., Olin, A. L., Murchoch, A. D., Canfield, A., Sasai, T., Timpl, R., Hardingham, T. E., and Aspberg, A. (2004) J. Biol. Chem. 279, 12511–12518
24. Xue, J. H., Nonoguchi, K., Fukumoto, M., Sato, T., Nishiyama, H., Higashitsuji, H., Itoh, K., and Fujita, J. (1999) Free Radic. Biol. Med. 27, 1238–1244
