The Cysteine-rich Protein Family of Highly Related LIM Domain Proteins*

(Received for publication, March 8, 1995, and in revised form, August 16, 1995)

Ralf Weiskirchen, Josephine D. Pino†, Teresita Macalma§, Klaus Bister‡**, and Mary C. Beckerle¶‡

From the Institute of Biochemistry, University of Innsbruck, Peter-Mayr-Straße 1a, A-6020 Innsbruck, Austria and the Department of Biology, University of Utah, Salt Lake City, Utah

Here we describe a family of closely related LIM domain proteins in avian cells. The LIM motif defines a zinc-binding domain that is found in a variety of transcriptional regulators, proto-oncogene products, and proteins associated with sites of cell-substratum contact. One type of LIM-domain protein, called the cysteine-rich protein (CRP), is characterized by the presence of two LIM domains linked to short glycine-rich repeats and a potential nuclear localization signal. We have identified and characterized two evolutionarily conserved members of the CRP family, CRP1 and CRP2, in chicken and quail. Expression of the genes encoding both CRP1 and CRP2 is differentially regulated in normal versus transformed cells, raising the possibility that members of the CRP family may function in control of cell growth and differentiation.

A number of genes are specifically and rapidly up-regulated in response to growth factor stimulation (1). The expression of these primary response, or immediate early, genes is independent of new protein synthesis and requires only the activation of pre-existing transcriptional regulators (2). The primary response genes encode proteins including transcription factors, proto-oncogene products, and other regulatory proteins that facilitate the transition of the cell from an arrested to a proliferative growth state and stimulate differentiation (3–5). The first cysteine-rich protein family member to be described (referred to here as CRP1)† was shown to be encoded by the primary response gene, CSR1, that exhibits serum induction (6, 7).

From analyses of both human and chicken cDNA, genomic DNA, and protein sequences (6–9), it has been determined that CRP1 contains two copies of a specific amino acid sequence motif termed LIM. The LIM motif displays the consensus amino acid sequence CX2CXC16–21H2CX2CX2CX16–23H2CX2–3(C/H/D) (8). Spectroscopic studies of LIM domains derived from a number of different proteins have revealed that the LIM domain specifically coordinates two zinc ions (10–13). The solution structure of a LIM domain derived from chicken CRP1 has been solved by two-dimensional NMR and illustrates that the LIM domain is itself a bipartite structure with spatially distinct modules focused around each metal binding site (14). Interestingly, although the LIM domain has been clearly demonstrated to function in specific protein-protein interactions (15–17), the tertiary fold of one zinc-binding module within the LIM domain is essentially identical to that found in well characterized DNA-binding zinc fingers (14). It remains to be determined whether the structural features of the LIM domain reflect a biologically significant ability to associate with nucleic acids as well as proteins.

The LIM motif was first identified in three developmentally regulated transcription factors, C. elegans lin-11, rat jls-1, and C. elegans mec-3, from which the term LIM is derived (10, 19). The LIM domain is often found in association with obvious functional domains, such as a DNA-binding homeodomain (18, 19) or a kinase domain (20). However, the LIM domain may also represent the primary sequence element in a protein. Examples of such "LIM only" proteins include CRP1 (6–9), the cysteine-rich intestinal protein (21), and rhombotin (22–25). Interestingly, although the LIM-only proteins lack DNA-binding homeodomains, they may also function in the regulation of cell growth and differentiation. For example, rhombotin-2 is a proto-oncogene product that is required for erythroid differentiation during mouse development (26) and overexpression of rhombotin genes in the thymus of transgenic mice results in T-cell acute lymphoblastic leukemia (27, 28).

CRP1 has been purified to homogeneity from chicken smooth muscle, and many of its biochemical and biophysical properties have been characterized (9). Binding studies have revealed that CRP1 interacts directly with another LIM protein called zyxin (8, 9). Both zyxin and CRP1 are localized at sites of membrane-substratum contact in association with the actin cytoskeleton (9). Together these proteins are postulated to perform a regulatory or signaling function at the adhesive membrane (8, 9, 15).

Recently, the level of a quail transcript that encodes a LIM domain protein with a high degree of structural similarity to CRP1 was shown to be dramatically reduced in avian fibroblasts transformed by retroviral oncogenes or chemical carcinogens (29). The amino acid sequence of the predicted quail protein was 79.8% identical to that of human CRP1 (29), and the protein was therefore postulated to represent the quail homologue of human CRP1. Subsequently, the complete amino acid sequence of chicken CRP1 was deduced from a cDNA clone and shown to be 90.6% identical to human CRP1 (9). Because...
the quail protein, although obviously closely related to CRP1, was significantly less similar to human CRP1 than was chicken CRP1. We postulated that the quail protein represented a new member of the CRP family, a CRP2. In this report, we demonstrate that both chicken and quail display multiple genes encoding CRP proteins. We have characterized two members of this gene family, CSRP1 and CSRP2, in both species.

**EXPERIMENTAL PROCEDURES**

**Cells and Viruses—**Chicken or quail embryo fibroblasts were prepared as described previously (30). The quail embryo fibroblast nonproducers (CSRP−) were transformed by the v-myc oncogene of avian myelocytomatosis virus MC29 and the quail embryo fibroblast producer line MH2-A2 transformed by the v-myc-v-mi oncogenes of the avian retrovirus MH2 have been described before (30, 31). The Rous sarcoma virus mutant tsLA29 encoding a temperature-sensitive v-src protein was a kind gift of J ohn A. Wyke (32). Infection of quail embryofibroblasts with MH2(MHAV) virus released from MH2-A2 cells and transformation with tsLA29 donor proviral DNA were carried out as described previously (30, 33).

Isolation and Characterization of Quail CSRP1 cDNA—Duplicate plaque lifts of a quail embryo fibroblast cDNA library prepared in gt10 (29) were screened by hybridization to quail CSRP2 cDNA (29) or to chicken CSRP1 cDNA (9) probes. The probes were prepared with the Multiprime DNA labeling system (Amersham Corp.). Hybridization was performed for 16 h at 37°C in SSPE (150 mM sodium chloride, 15 mM sodium citrate), 5 × Denhardt's solution (0.1% [w/v] Ficoll, 0.1% [w/v] bovine serum albumin, 0.1% [w/v] polyvinylpyrrolidone), 5 mM EDTA, 0.5% (w/v) SDS, and 50 μg/ml sheared, denatured salmon sperm DNA. Following hybridization, the filters were washed once for 20 min at 55°C on a rotator, 1× EDTA, and 0.1% (w/v) SDS and then as follows: 20 min at 50°C in 0.4× SSC, 1 mM EDTA, and 0.1% (w/v) SDS. A screen of 3 × 106 recombinant plasmid phage led to the identification of about 210 phage positive for both probes. Comparison of the hybridization signals obtained with the quail CSRP2 and chicken CSRP1 probes revealed that about 40 of these recombinant phage harbored cDNA inserts. Two of these phage were purified using standard procedures, and their cDNA inserts were subcloned into the EcoRI site of the plasmid vector pUC19. Nucleotide sequences of plasmid DNAs were directly determined by the dideoxynucleotide chain termination method, using a T7 sequencing kit (Pharmacia Biotech Inc.). Both clones were positively identified by partial sequence comparison with the reported sequence of chicken CSRP1 cDNA (9). The complete nucleotide sequence of one clone (clone 7) was determined by sequencing of the two internal BglI and the unique internal PvuI restriction sites.

Isolation and Characterization of Chicken CSRP2 cDNA—A chicken embryo cDNA library cloned into the EcoRI site of the expression vector gt11 (Clontech Laboratories Inc., Palo Alto, CA) was screened by hybridization to chicken CSRP1 cDNA (9) or to quail CSRP2 cDNA (29). Duplicate plaque lifts were probed with both chicken CSRP1 probe and quail CSRP2 probe. Both the chicken CSRP1 and quail CSRP2 probes were labeled with 32P by the random primer method according to an optimized method (34) and Sequenase I (U.S. Biochemical Corp.), or by dideoxynucleotide chain termination sequencing of polymerase chain reaction products (Life Technologies, Inc.). Ambiguities were resolved using a modification of the dideoxynucleotide chain termination method (35). Chicken genomic DNA containing a CSRP1 cDNA probe was isolated by the polymerase chain reaction using primers corresponding to nucleotides 354–374 and 558–577 of CSRP2-TM1 cDNA. Sequencing of the genomic DNA fragment was conducted using methods described above.

Identification of the 5′ End of Chicken CSRP2 cDNA—The 5′ end of chicken CSRP2 was identified using two methods. First, sequences corresponding to the 5′ end of CSRP2 mRNA (referred to as CSRP2 probes) were isolated using the RACE procedure (Clontech Laboratories, Inc., Palo Alto, CA) (36). The poly(A) + mRNA utilized in this procedure was purified from approximately 2 × 107 chicken embryo fibroblasts using a commercially available protocol (Qiagen Inc., Chatsworth, CA). All primers were synthesized using an Applied Biosystems DNA synthesizer, model 380 B. In addition, a chicken embryo fibroblast agt11 cDNA library (37) was probed using methods described above. Four clones containing the 5′ end of chicken CSRP2 were identified and analyzed by DNA sequencing, confirming the results obtained with the 5′ RACE procedure.

**Southern Blot Analysis of Quail and Chicken Genomic DNA—**High molecular weight quail genomic DNA was isolated from quail embryo fibroblasts following a proteinase K/phenol extraction protocol (31). Equal amounts of quail genomic DNA or chicken genomic DNA (Promega, Madison, WI) were digested with EcoRI, BamHI, or PstI. The digests were separated according to size by electrophoresis through agarose (10 μg/ml) and transferred to Hybond-N nylon membrane (Amersham Corp.) using standard techniques (38). Sequences corresponding to the coding regions of quail and chicken CSRP1 and CSRP2 cDNAs were labeled with [32P]dCTP as described above and used as hybridization probes. Hybridizations were performed for approximately 16 h at 65°C in 5× SSPE, 5 × Denhardt's solution, 0.5% (w/v) SDS, and 0.1% (w/v) SDS. Following hybridization, the filters were washed twice for 10 min at room temperature in a solution containing 2× SSPE and 0.1% (w/v) SDS, then for 15 min at 65°C in 1× SSPE and 0.1% SDS and then for 10 min at 65°C in 0.1× SSPE and 0.1% (w/v) SDS.

**Northern Blot Analysis of Quail and Chicken RNA—**Preparation of total cellular RNA and selection of poly(A) + RNA were performed essentially as described previously (29). For northern analyses, 7-μg portions of poly(A) + RNA or 30-μg portions of total RNA from normal or transformed chicken or quail embryo fibroblasts were separated by electrophoresis on 1% (w/v) agarose, 2.2% formaldehyde gels as described (39). Hybridizations were carried out for 16 h at 37°C in 35 ml of buffer H using 32P-labeled insert DNAs from chicken or quail CSRP1 and CSRP2 DNA clones or from a quail GAPDH cDNA clone (39) as probes. Hybridized filters were washed twice at 55°C in 2× SSC, 1× EDTA, and 0.1% (w/v) SDS and then twice for 20 min at 50°C in 0.4× SSC, 1× EDTA, and 0.1% (w/v) SDS.

Characterization of the Two Chicken CSRP1 Transcripts—Two CSRP1 transcripts are observed by Northern analysis. The originally reported chicken CSRP1 cDNA sequence corresponds to the longer transcript, CSRP1.1 (9). Reexamination of the collection of CSRP1 cDNAs isolated in the initial library screen revealed a cDNA that corresponded to the shorter transcript, CSRP1.2. This transcript was subcloned into the EcoRI restriction site of pBS to generate pBS-CSRP1.2. The sequence of CSRP1.2 cDNA was determined on both strands as described above. The sequences of the 3′-untranslated regions of CSRP1.1 and CSRP1.2 were compared using the program “Bestfit” (40).

**Comparisons of Amino Acid Sequences—**Alignments of pairs of amino acid sequences were performed by the method of Myers and Miller (41), and multiple sequence alignments were performed by the method of Higgins and Sharp (42). Both algorithms were used as parts of the PC/Genie program package (Intelligenetics Inc., Mountain View, CA).

Nomenclature—The first CRP family member to be identified (6–9) is CSRP1. The genes encoding the CRPs are referred to with the symbol CSRP1, CSRP2, etc. It was necessary to use the symbol CSRP1 because the gene designation CSRP1 has already been assigned to an unrelated gene. The nomenclature used to describe avian proteins and genes in this report conforms to that first developed for the human proteins and genes (43).
RESULTS

As described above, previous work had suggested that multiple CSRP family members might be represented in the vertebrate genome. Here we have explored the possibility that two avian species, the chicken and the quail, express multiple forms of CRP. In this paper, the cysteine-rich protein that was originally described in human (6, 7) and chicken (8, 9) is referred to as CRP1. The closely related gene product that was originally described in quail (29) is here referred to as CRP2 to indicate that we consider it to be a member of the same family. The CRSP gene symbols refer to genes that encode members of the CRP family of proteins.

Identification and Characterization of Quail CSRP1 cDNA—CSRP1 had been identified unequivocally in humans (6, 7) and chickens (8, 9), but not yet in quail, in which the CSRP2 gene was originally identified (29). A chicken CSRP1 cDNA probe (9) was used to screen a quail cDNA library in an effort to identify the quail sequences that are most closely related to CSRP1. 3 × 10^6 recombinants were screened, and about 40 clones carrying sequences that hybridized strongly with the chicken CSRP1 probe were identified. A 927-base pair cDNA clone (clone 7) was isolated, and its complete nucleotide sequence was determined (Fig. 1). The cDNA clone (clone 7) was isolated, and its complete nucleotide sequence was determined (Fig. 1). The conserved cysteine and histidine residues that define the two LIM domains are circled. The sequence is deposited in the EMBL data base under the accession number Z28333.

Identification and Characterization of Chicken CSRP2 cDNA—A quail CSRP2 cDNA probe (29) was used to screen a chicken embryo cDNA library in an effort to identify sequences that are most closely related to CSRP2. 1.5 × 10^6 recombinants were screened, and 26 strongly reactive clones were identified and isolated. The cDNA inserts carried in seven of the initial phage isolates were characterized further; two are described in detail here. One clone referred to as CSRP2-TM1 contained a 783-base pair insert with an open reading frame that encodes 192 amino acids but lacks a translation initiation codon (Fig. 2A). Another clone isolated in this initial screen, designated CSRP2-TM6, contained a deletion of two nucleotides corresponding to positions 530 and 531 in the composite CSRP2 sequence shown in Fig. 2B. This deletion resulted in the generation of an in-frame stop codon (TGA) that disrupted the open reading frame of CSRP2 within the second LIM domain, causing the elimination of 27 C-terminal amino acid residues. None of the other clones that were analyzed exhibited this two-base pair deletion. Nevertheless, in order to evaluate the possible significance of this sequence heterogeneity, we isolated and sequenced chicken genomic DNA in this region and

![Fig. 1. Nucleotide sequence of quail CSRP1 cDNA and deduced amino acid sequence of the quail CRP1 protein.](image)

found that the chicken genome contained sequences that were compatible with the CSRP2-TM1 cDNA, but not the CSRP2-TM6 cDNA. Consequently, we conclude that the CSRP2-TM1 cDNA accurately reflects the sequence information present in the genome, and the two-base pair deletion observed in CSRP2-TM6 is likely to reflect an artifact introduced during the generation or amplification of the cDNA library.

In multiple CSRP2 cDNAs that were characterized, none extended the 3' end beyond what was observed for CSRP2-TM1. Likewise, none of the clones isolated in the original screen contained any useful 5' sequence beyond what was found in CSRP2-TM1. Therefore, in order to identify cDNA clones that contained the 5' end of the coding sequence and the
5′-untranslated region corresponding to the CSRP2 mRNA, we used two strategies. First, we employed a modified primer extension-5′-RACE technique (36) to identify the 5′ end of the CSRP2 cDNA; a representative cDNA clone derived from this screen is referred to as CSRP2-5′-RACE (Fig. 2A). In addition, we screened an independently generated chicken embryo fibroblast cDNA library (37) with the quail CSRP2 probe. Four clones that were identified in this screen displayed 5′ extensions of the cDNA insert that contained the remainder of the coding sequence as well as some of the 5′ untranslated region as observed in the CSRP2-5′-RACE cDNA.

The nucleotide and deduced amino acid sequences of the composite CSRP2 cDNA derived from the fusion of CSRP2-TM1 with the product of the 5′-RACE is shown in Fig. 2B. The initiation codon, termination codon, and polyadenylation signal are underlined. The predicted chicken CRP2 protein is 194 amino acids in length with a calculated molecular weight of 20,925 and a predicted unmodified pl of 8.68. As in the case of other CRP family members, chicken CRP2 displays two LIM domains with associated glycine-rich motifs and a potential nuclear localization signal. The metal-coordinating cysteine and histidine residues that contribute to the LIM consensus are boxed in Fig. 2B.

The chicken CRP2 amino acid sequence is 76.6% identical to chicken CRP1 and 99.5% identical to quail CRP2, with only a single conservative amino acid substitution at residue 95 distinguishing the chicken and quail CRP2 homologues.

Expression of CSRP1 and CSRP2 in Normal and Transformed Avian Fibroblasts—CSRP2 was originally identified in quail cells by a differential screen for genes whose expression was altered in quail embryo fibroblasts that were transformed with the retroviral oncogene v-myc (29). Subsequent analysis revealed that CSRP2 expression was dramatically reduced in quail cells that were transformed with a variety of oncogenes including v-myc, v-src, and v-myc-v-mil in cells that were derived from a methylcholanthrene-induced quail fibrosarcoma (29). These studies raised the possibility that the expression of a CSRP family member could be important for the regulation of controlled cell growth and differentiation.

We have extended these studies here to evaluate whether CSRP1 and CSRP2 behave in a similar manner when cells are challenged with a transforming factor. The levels of CSRP1 and CSRP2 mRNAs in normal chicken embryo fibroblasts, normal quail embryo fibroblasts, and transformed quail embryo fibroblasts were evaluated by Northern analysis (Fig. 3). In chicken embryo fibroblasts, the CSRP1 probe recognizes two transcripts, a minor species of 1.4 kb and a major species of 1.0 kb, but only a single transcript of 1.0 kb is detected in quail embryo fibroblasts (Fig. 3, A and C). The CSRP2 probe hybridizes to a single transcript of 0.9 kb that is present in both chicken and quail embryo fibroblasts (Fig. 3, B and D). CSRP1 expression is strongly decreased in the v-myc-transformed quail embryo fibroblast line Q8 (Fig. 3, A and C). However, longer exposures of the autoradiographs (not shown) reveal the presence of residual CSRP1 mRNA in transformed cells. In contrast, there is a nearly complete loss of CSRP2 transcripts in such cells (Fig. 3, B and D), in confirmation of our previous results. The level of mRNA encoding the enzyme glyceraldehyde-3-phosphate dehydrogenase was used as an internal control for the quality and amount of mRNA present in each lane (39).

In order to rule out the possibility that the suppression of CSRP1 and CSRP2 expression occurs only in long term transformed cell lines such as Q8, we analyzed their expression in a conditional transformation system and during the process of the initial establishment of oncogene-induced transformation.

Fig. 2. Nucleotide sequence of chicken CSRP2 cDNA and deduced amino acid sequence of the chicken CRP2 protein. A, a schematic representation of the isolated CSRP2 cDNA clones used to generate the composite CSRP2 cDNA; B, the nucleotide (numbered on the right) and deduced amino acid (numbered on the left) sequences of the composite cDNA clone (CSRP2) are shown. The translational initiation codon, stop codon, and the polyadenylation signal are underlined. Cysteine and histidine residues that contribute to the LIM consensus motif are boxed. Glycine residues that are found in the repeat adjacent to each LIM domain are circled. The sequence is deposited in the EMBL data base under the accession number X84264.
signals would be predicted to give rise to mRNA transcripts of sequence at nucleotides 942–947 and a CACTG recognition element at nucleotides 1312–1316. Utilization of both of these signals would be predicted to give rise to mRNA transcripts of approximately 1.4 and 1.0 kb, consistent with what is observed by Northern analysis (Fig. 3). Characterization of a number of CSRP1 cDNA clones isolated from a chicken embryo fibroblast cDNA library revealed that two classes of cDNAs were present. The two classes were indistinguishable in their coding sequences and varied only in their 3′-untranslated regions, in particular with respect to the position of the poly(A) tail (Fig. 5). In the large sized minor transcript, referred to as CSRP1.1 mRNA, the CACTG site appears adjacent to the poly(A) tail in a 3′-untranslated region of 668 nucleotides, whereas the small sized major transcript, referred to as CSRP1.2 mRNA, appears to result from the alternative use of the AATAAA signal to generate a 3′-untranslated region of only 310 nucleotides excluding the poly(A) tail.

The Complexity of the CSRP Gene Family—Southern blot analysis of the following cellular sources is shown: quail embryo fibroblasts (lane 1); quail embryo fibroblasts transformed by the temperature-sensitive protein product of the v-src oncogene of tsLA29 and kept at 37°C (lane 2) or shifted to 40.5°C for 1 day (lane 3) or 2 days (lane 4); quail embryo fibroblasts transformed by tsLA29 at 37°C and then shifted and kept at 40.5°C (lane 5) or shifted back to 37°C for 1 day (lane 6) or 2 days (lane 7). B, a Northern blot analysis of total RNAs (30 μg/ lane) from the following cellular sources is shown: quail embryonic fibroblasts (lane 1); quail embryo fibroblasts at day 7 (lane 2), day 16 (lane 3) or day 29 (lane 4) postinfection with the avian retrovirus MH2 carrying the two oncogenes v-myc and v-mil; the MH2-transformed quail fibroblast line MH2-A10 (lane 5). RNAs from both filters (A and B) were first hybridized with a 32P-labeled quail CSRP2 probe (3.4 × 10⁶ cpm), and the autoradiograph was exposed for 7.5 h using an intensifying screen. The filters were stripped and then hybridized with a 32P-labeled quail CSRP1 probe (6.7 × 10⁶ cpm), and the autoradiograph was exposed for 15 h. The positions of ribosomal RNAs are indicated in the margin. Hybridization with a 32P-labeled glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probe (4.9 × 10⁶ cpm; 8-h exposure) was used as an internal control.
analysis was used to confirm that CSRP1 and CSRP2 sequences are derived from distinct genes. When chicken genomic DNA was digested with BamHI, EcoRI, or PstI and probed at high stringency with either a chicken CSRP1 probe (Fig. 6A) or a CSRP2 probe (Fig. 6B), a simple pattern of nonoverlapping bands was detected. When the same blots were probed at lower stringency (not shown), substantial cross-hybridization of CSRP1 and CSRP2 sequences was observed, indicating that CSRP1 and CSRP2 represent distinct but related genes in the chicken. Similar experiments using quail genomic DNA confirmed that CSRP1 and CSRP2 are distinct genes that display cross-hybridization under low stringency conditions (Fig. 5C and D, data not shown).

Interestingly, low stringency Southern blots also revealed the presence of some minor cross-hybridizing DNA fragments that were not detected in the high stringency screens with either the CSRP1 or CSRP2 probes (not shown). These bands may represent sequences derived from more distantly related CSRP family members. Indeed, a recent report describes an- other avian gene, CSRP1, adjacent glycine-rich repeats, and a potential nuclear targeting signal (KKYGPK) is found at the same position (amino acid residues 64–69) in all CRP proteins. Its sequence is almost completely conserved, with only one chemical conservative substitution of an arginine for a lysine at position 64 in the case of MLP/CRP3. The distance between the two LIM domains is also highly conserved in all six CRP family members display exactly 17 amino acid residues between the second and third, and between the sixth and seventh, metal-coordinating residues. In contrast, other LIM proteins may display anywhere from 16 to 23 residues in these positions (8). Moreover, in CRP family members, each LIM domain is proximal to a glycine-rich repeat that closely conforms to the sequence originally found in human CRP1 (6). A potential nuclear targeting signal (KKYGPK) is found at the same position (amino acid residues 64–69) in all CRP proteins. Its sequence is almost completely conserved, with only one chemical conservative substitution of an arginine for a lysine at position 64 in the case of MLP/CRP3. The distance between the two LIM domains is also highly conserved in all six CRP proteins, varying only from 56 to 59 residues. This conservation is the relative positions of the LIM domains is particularly striking given the heterogeneous spacing between LIM domains found in other proteins (Fig. 8). For example, within a single protein, two LIM domains may be separated by as few as 7 (8) or as many as 68 (46) amino acids. Finally, the sequence conservation of the CRP family members is high (Fig. 7, Table I). Human, chicken, and quail CRP1s share more than 90% amino acid sequence identity. The CRP2 sequences are the next most closely related sequences to CRP1, exhibiting 76–80% identity to the CRP1 sequences. MLP/CRP3 retains between 63 and 67% sequence identity to CRP1 and CRP2 proteins. Collectively, the common features described above

Fig. 5. Differential polyadenylation of chicken CSRP1 transcripts. The two chicken CSRP1 transcripts observed by Northern analysis (cf. Fig. 3, A and C) result from alternative polyadenylation. The 3′-untranslated regions of the longer transcript (CSRP1) and the shorter transcript (CSRP1.2) are shown. The sequences corresponding to polyadenylation recognition signals (AATAAA and CACTG) are underlined.

Fig. 6. Chicken and quail CSRP1 and CSRP2 genes represent distinct genetic loci. Chicken (A and B) and quail (C and D) genomic DNA (10 µg/lane) were digested with BamHI, EcoRI, or PstI, and the digests were analyzed by Southern blotting using 32P-labeled chicken CSRP1 (A), chicken CSRP2 (B), quail CSRP1 (C), or quail CSRP2 (D) cDNA clones as hybridization probes. The positions of DNA size markers are indicated in the margin.
illustrate that the architecture of CRP family members is conserved and is distinct from that of other LIM proteins.

**DISCUSSION**

The CRP Family of LIM Only Proteins—The discoveries of quail (29) and chicken (8, 9) genes encoding cysteine-rich LIM-domain proteins (CRPs) that were closely related to each other and yet showed significantly different degrees of homology to a human protein of similar structure (6, 7) had raised the possibility that the two avian genes were not homologous counterparts of each other but rather represented closely related members of a multigene family. The results presented in this paper provide direct proof for this suggestion. For each of the two genes originally identified, quail CRP1 and chicken CRP1 (Fig. 1), human CRP1 (6), chicken CRP2 (Fig. 2), quail CRP2 (29), and chicken MLP/CRP3 (45) is shown. Residues conserved in all sequences are marked by asterisks below the bottom line. The conserved Cys and His residues of the LIM domains are indicated by crosses, and the conserved glycine residues of the glycine-rich repeats are marked by dashes above the top line.

Comparisons of a number of LIM-proteins (Fig. 8) reveal the conserved structural elements of the CRP family members that set them apart from other LIM only proteins. Both the abundance and spacing of the metal-coordinating residues (cysteine and histidine) are absolutely conserved in all individual LIM domains of all CRP-type proteins identified so far. Furthermore, the spacing between the two LIM domains is nearly identical in all CRP proteins, and the location and sequence of glycine-rich repeats and potential nuclear localization signals are shared among all of these proteins. In addition to this highly conserved domain topography, the overall sequence identities between different types of CRPs are further indications of the close evolutionary relationships of these proteins (Fig. 7, Table I). We therefore conclude that avian and mammalian species contain in their genomes a multigene family specifying different but very closely related CRP proteins. We suggest a uniform nomenclature for these proteins, CRP1 and CRP2 for the proteins described here and previously (6–9, 29) and CRP3 for the recently described muscle-specific LIM protein, MLP (45). It should be noted that a recently described rat LIM protein that is also referred to as a cysteine-rich protein (46) exhibits a general domain topography that is similar to the CRP family members. However, both the differences in spacing within and between the two LIM domains and low overall sequence identities distinguish the rat protein from members of the CRP family.

Biochemical and Biological Functions of CRPs—Given the prominence of LIM domains in CRPs (Fig. 8), it is likely that the biochemical functions of CRP family members will largely be defined by their tandem LIM domains. The LIM domain has

2 R. Weiskirchen and K. Bister, unpublished results.
been postulated to interact with both nucleic acids and proteins. Because a number of zinc-binding proteins are well-defined transcription factors, the LIM domain was originally thought to function in DNA binding. Interestingly, the analysis of a chicken CRP1 LIM domain by two-dimensional NMR spectroscopy revealed that the tertiary fold of one zinc-binding module within the CRP LIM domain is essentially identical to that determined for well characterized DNA-binding zinc fingers (14). These structural studies illustrated that the tertiary fold of the LIM domain may be compatible with nucleic acid binding; however, a specific association between a LIM domain and nucleic acid has yet to be demonstrated.

On the other hand, there is growing evidence that LIM domains mediate specific protein-protein interactions. Direct evidence for a functional role of LIM domains in protein-protein interactions was recently provided for zyxin and CRP1, two interacting chicken LIM-proteins (8, 9, 15). In addition, a LIM protein has been shown to interact with tyrosine-containing tight turn motifs present in the cytoplasmic domain of the insulin receptor (17). LIM domains have also been implicated in homotypic, intermolecular interactions (16). Moreover, interactions of LIM domain proteins with proteins containing basic helix-loop-helix motifs known to be involved in protein dimerization have been demonstrated (47, 48). Thus, the LIM domain can clearly support specific associations with partner proteins. Proteins that display multiple LIM domains may serve as adaptor molecules or as scaffolds for the coordinated, localized assembly of multimeric protein complexes (15).

The detailed biological role of the CRP family of proteins is not well understood. In general, LIM domain proteins, in particular those that contain additional homeodomains, have been implicated in regulatory processes important for development and cellular differentiation (18, 19, 49). Many of the LIM-only proteins also appear to function in these broad processes. For example, the rhombotins were originally identified at chromosomal translocations and shown to be involved in tumorigenesis (22–24, 27, 28), and recent studies on their tissue-specific expression and in vivo function have revealed that rhombotins have essential roles in normal development (25, 26). For the CRP protein family, a proposed role in regulatory processes was most clearly confirmed for the MLP protein that was isolated from a subtracted cDNA library enriched in genes induced in skeletal muscle by denervation and then shown to be a positive regulator of myogenesis (45). For chicken CRP1, direct interaction with zyxin, an adhesion plaque protein, has been demonstrated, and it was postulated that both proteins may function as components of a signal transduction pathway that mediates adhesion-stimulated changes in gene expression (8, 9). The expression of the human CSRP1 gene was shown to be induced as a primary response to serum in quiescent cells, with a serum induction profile similar to that of c-myc, and expression that continues, like that of c-myc, in logarithmically growing cells (7). This is in agreement with our results reported here and previously (29) on CSRP1 and CSRP2 expression in continuously growing normal avian fibroblasts. The strong suppression of CSRP genes, in particular of CSRP2, in all transformed cells tested may be connected with a regulatory function of CRP proteins in ordered cell growth. Although there is strong circumstantial evidence for the involvement of CRP

![Fig. 8. Domain structure of selected LIM-proteins. A, a schematic diagram of the structures of representative chicken (c) or human (h) LIM proteins including members of the CRP family is shown. LIM motifs and glycine-rich repeats are shown as open and black boxes, respectively. The amino- and carboxyl-terminal amino acid residues and the first and last residues of the LIM domains are numbered. Sources for the amino acid sequences are as follows: chicken CRP1 (9), chicken CRP2 (Fig. 2), chicken MLP/CRP3 (45), human TTG1 (22), chicken zyxin (8). B, the spacing of the eight metal-coordinating amino acid residues (shown in boldface letters) within all LIM domains of the proteins shown in panel A is compared. Among the members of the CRP family (CRP1, CRP2, MLP/CRP3), the spacing is absolutely conserved. Numbering of LIM domains refers to their order of appearance in these proteins relative to the amino terminus and does not reflect a structural classification.](https://example.com/fig8.png)

### Table I

| Protein       | hCRP1 | cCRP1 | qCRP1 | cCRP2 | qCRP2 | dMLP/CRP3 |
|---------------|-------|-------|-------|-------|-------|-----------|
| hCRP1         | 100   | 90.6 (3.7) | 90.6 (3.7) | 79.8 (8.8) | 79.8 (8.8) | 66.3 (10.4) |
| cCRP1         | 90.6 (3.7) | 100 | 100 | 76.6 (9.9) | 76.6 (9.9) | 63.5 (9.9) |
| qCRP1         | 90.6 (3.7) | 100 | 100 | 76.6 (9.9) | 76.6 (9.9) | 63.5 (9.9) |
| cCRP2         | 79.8 (8.8) | 76.6 (9.9) | 76.6 (9.9) | 100 | 99.5 (0.5) | 100 |
| qCRP2         | 79.8 (8.8) | 76.6 (9.9) | 76.6 (9.9) | 99.5 (0.5) | 100 | 66.5 (10.8) |
| dMLP/CRP3     | 66.3 (10.4) | 63.5 (9.9) | 63.5 (9.9) | 66 (11.3) | 66.5 (10.8) | 100 |
proteins in regulatory processes important for cell growth and differentiation, definitive characterization of their biological functions and distinction between the functions of the individual members of the CRP family of closely related LIM only proteins will depend on the elucidation of their biochemical functions and on the identification of their cellular targets.

Why Have So Many CRPs?—The existence of multiple, closely related members of the CSRP gene family may reflect an organism’s need for functionally distinct CRP proteins or may be necessary for tissue-specific gene expression. It is often difficult to distinguish between these two possibilities, which may be necessary for tissue-specific gene expression. It is often difficult to distinguish between these two possibilities, which are not mutually exclusive. It will be interesting to determine whether the individual members of the CRP family are functionally distinct or if a collection of CSRP genes has evolved to allow for tissue-specific or temporally regulated expression of CRP isoforms with equivalent biochemical functions.

Acknowledgments—We appreciate the participation of Todd Alder in the early stages of this project and the excellent technical assistance by Salaine Waerichsen. The Human, Mouse, and Chicken Gene Nomenclature Committees have approved the use of the symbols CSRP (human and chicken) and Csrp (mouse) for the sequences encoding members of the CRP family of proteins.

REFERENCES

1. Bravo, R. (1990) Cell Growth & Differ. 1, 305–309
2. Lau, L. F., and Nathans, D. (1985) EMBO J. 4, 3145–3151
3. Kelly, K., Cochran, B. H., Stiles, C. D., and Leder, P. (1983) Cell 35, 603–610
4. Müller, R., Bravo, R., Burkhardt, J., and Curran, T. (1994) Nature 321, 716–720
5. Greenberg, M. E., and Ziff, E. B. (1984) Nature 311, 433–438
6. Liebhaber, S. A., Emery, J. G., Urbanek, M., Wang, X., and Cooke, N. E. (1990) Nuclet Acids Res. 18, 3871–3879
7. Wang, X., Lee, G., Liebhaber, S. A., and Cooke, N. E. (1992) J. Biol. Chem. 267, 9176–9184
8. Sadler, J., Crawford, A. W., Michelsen, J. W., and Beckerle, M. C. (1992) J. Cell Biol. 119, 1573–1587
9. Crawford, A. W., Pino, J. D., and Beckerle, M. C. (1994) J. Cell Biol. 124, 117–127
10. Michelsen, J. W., Schmeichel, K. L., Beckerle, M. C., and Wing, D. R. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 4404–4408
11. Michelsen, J. W., Sewell, A. K., Louis, H. A., Olsen, J. I., Davis, D. R., Winge, D. R., and Beckerle, M. C. (1994) J. Biol. Chem. 269, 11108–11113
12. Koss, J. J., Michelsen, J. W., Louis, H. A., Olsen, J. I., Davis, D. R., Beckerle, M. C., and Wing, D. R. (1994) Biochemistry 33, 468–477
13. Archer, V. E. V., Breton, J., Sanchez-Garcia, I., Osada, H., Forster, A., Thomson, A. J., and Rabbitts, T. H. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 316–320
14. Perez-Alvarado, G. C., Miles, C., Michelsen, J. W., Louis, H. A., Wing, D. R., Beckerle, M. C., and Summers, M. F. (1994) Nature Struct. Biol. 1, 388–398
15. Schmeichel, K. L., and Beckerle, M. C. (1994) Cell 79, 211–219
16. Feuerstein, R., Wang, X., Song, D., Cooke, N. E., and Liebhaber, S. A. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 10655–10659
17. Wang, Y.-Y., and Gill, G. N. (1994) J. Biol. Chem. 269, 25085–25090
18. Fried, G., Kim, S. K., and Horvitz, H. R. (1990) Nature 344, 876–879
19. Karlsson, O., Thor, S., Norberg, T., Ohlsson, H., and Edlund, T. (1990) Nature 344, 879–882
20. Mizuno, K., Okano, I., Ohashi, K., Nounoue, K., Kuma, K., Miyata, T., and Nakamura, T. (1994) Oncogene 9, 1605–1612
21. Birkenmeier, E. H., and Gordon, J. I. (1996) Proc. Natl. Acad. Sci. U. S. A. 83, 2516–2520
22. McGuire, E. A., Hockett, R. D., Pollock, K. M., Bartholdi, M. F., O’Brien, S. J., and Korsmeyer, S. J. (1989) Mol. Cell. Biol. 9, 2124–2132
23. Boehm, T., Foroni, L., Kaneko, Y., Perutz, M. F., and Rabbitts, T. H. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 4367–4371
24. Royer-Pokora, B., Loos, U., and Ludwig, W.-D. (1991) Oncogene 6, 1887–1893
25. Foroni, L., Boehm, T., White, L., Forster, A., Sherrington, P., Liao, X. B., Brannan, C. I., Jenkins, N. A., Copeland, N. G., and Rabbitts, T. H. (1992) J. Mol. Biol. 226, 747–761
26. Warren, A. J., Colledge, W. H., Carlton, M. B. L., Evans, M. J. J., Smith, A. H. J., and Rabbitts, T. H. (1994) Cell 78, 45–57
27. McGuire, E. A., Rintoul, C. E., Sdor, G. M., and Korsmeyer, S. J. (1992) Mol. Cell. Biol. 12, 4186–4196
28. Fish, P., Boehm, T., Lavener, J., Larson, T., Arno, J., Forster, A., and Rabbitts, T. H. (1992) Oncogene 7, 2389–2397
29. Weiskirchen, R., and Bister, K. (1993) Oncogene 8, 2317–2324
30. Bister, K., Hayman, M. J., and Vogt, P. K. (1977) Virology 82, 431–448
31. Jessen, H. W., Patschinsky, T., and Bister, K. (1983) J. Virol. 46, 61–73
32. Welham, M. J., and Wyke, J. A. (1988) J. Virol. 62, 1898–1906
33. Hartl, M., Vogt, P. K., and Bister, K. (1995) Virology 207, 321–326
34. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
35. Schuurman, R., and Keulen, W. (1991) BioTechniques 10, 185
36. Edwards, J. B. D. M., Delort, J., and Mallett, J. (1991) Nucleic Acids Res. 19, 5227–5232
37. Tamkun, J. W., DeSimone, D. W., Fonda, D., Patel, R. S., Buck, C., Horvitz, A. F., and Hynes, R. O. (1986) Cell 46, 271–282
38. Schmeichel, K. L., and Beckerle, M. C. (1994) Trends Genet. 10, 315–320