

Avian cellular homolog of the *qin* oncogene

HWAI WEN CHANG*, JIAN LI*, DORIS KRETZSCHMAR†, AND PETER K. VOGT

The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037

Contributed by Peter K. Vogt, October 7, 1994

ABSTRACT We have isolated chicken cDNA clones of the *c-qin* gene, the cellular counterpart of the *v-qin* (Chinese for “avian”) oncogene of avian sarcoma virus 31. There are several differences between the cellular and the viral *qin* sequences: (i) two nonconservative amino acid substitutions in the *Qin* coding region; (ii) a truncation in the carboxyl terminus of the viral protein due to a premature stop codon; (iii) a partial *Gag* sequence fused to the amino terminus of viral *Qin*; and (iv) eight cell-coded amino acids which link the cellular *Qin* coding domain to the viral *Gag* domain. We have also characterized the expression pattern of *c-qin* in chicken embryos by *in situ* hybridization and by Northern blot analysis. *c-qin* is abundantly expressed in the developing brain, and this expression is restricted to the telencephalon of early embryos.

The *v-qin* oncogene is a cell-derived insert in the genome of avian sarcoma virus 31 and functions as the oncogenic determinant of that virus. The virus causes fibrosarcomas in chickens and transforms chicken embryo fibroblasts (CEFs) *in vitro* (1). The *v-Qin* protein contains a region with close homology to the conserved DNA-binding domain of the winged helix family of transcription factors, also referred to as the hepatocyte nuclear factor 3 (HNF3)/fork head (Fkh) family (2, 3). Winged helix proteins are conserved in metazoans and function as regulators of embryonal development and tissue differentiation. Members of this family are related mainly by a 110-aa DNA-binding domain that consists of three α -helices and β -sheets grouped in a tight arrangement, forming a variant of the helix-loop-helix motif, referred to as a winged helix (4). Proteins of the winged helix family are expressed in a highly tissue-specific manner, consistent with their role in development and differentiation. The Fkh protein of *Drosophila*, a prototype member of the family, is required for the proper formation of the terminal structures of embryos (5). Mutation of the *Drosophila fkh* gene causes replacement of the ectodermal portions of the gut by ectopic head structures at both ends of the embryo. In *Drosophila*, seven *fkh*-related genes have so far been identified; of these, sloppy paired 1 (*slp-1*) and sloppy paired 2 (*slp-2*) are essential for normal morphogenesis (6). In vertebrates, the winged helix family includes numerous developmentally regulated proteins (7). Examples are XFD-1, the *Xenopus laevis* homolog of the Fkh protein, which is activin-inducible (8); rat brain factor 1 (BF-1), which is expressed only in the telencephalon region of the neural tube (9); and QRF-1, which is restricted to B cells and muscle (10). The rat liver transcription factors HNF-3 α , -3 β , and -3 γ , together with several other factors that are enriched in the liver, regulate the expression of genes important for liver function (3, 11, 12).

The function of the *qin* gene is unknown, but its inclusion in the winged helix family suggests that it may encode a transcriptional regulator. The ability of *v-qin* to induce neoplastic transformation further suggests a role in the regulation of cell growth. Within the winged helix family, the *Qin* protein is most closely related to the rat BF-1 protein and may be the avian homolog of BF-1 (9). The carboxyl-terminal region of the *v-Qin* protein (aa 124–387) shows 97.7% homology with BF-1,

but the amino-terminal portion of the *v-Qin* protein diverges significantly from BF-1. This paper describes the isolation and sequencing of cDNA clones of chicken *c-qin*.[‡] Northern blot analysis of tissues from chicken embryo shows that *c-qin* RNA is expressed most abundantly in the forebrain. *In situ* hybridizations of whole-mount chicken embryos and of sections demonstrate that the expression of *c-qin* is restricted to the telencephalon of developing chicken embryos.

MATERIALS AND METHODS

Isolation and Sequencing of cDNA Clones of Chicken *c-qin*. Chicken *c-qin* cDNA was obtained by screening a λ gt11 chicken embryonic cDNA library (provided by J. Sap, European Molecular Biology Laboratory) with a *v-qin* probe which extends from nt 768 to nt 1397 (1). This probe does not include sequences that code for the conserved winged helix DNA-binding domain. Hybridization was performed in buffer containing 50% formamide, 5 \times standard saline citrate (SSC), 20 mM Na₂HPO₄/NaH₂PO₄ (pH 7.6), 7% SDS, 1% PEG (*M_w* 20,000), and 0.5% nonfat powdered milk. Final washes were in 1 \times SSC/0.1% SDS at 60°C. From 400,000 clones, three positive clones were obtained. The largest clone, CEQ 3-1, was used for further studies. The CEQ 3-1 cDNA was sequenced in both strands with the Sequenase version 2.0 kit (United States Biochemical).

Northern Blot Analysis of Chicken Embryo Tissues. Poly(A)⁺ RNA was isolated from 18-day chicken embryo tissues and from adult chickens with the mRNA Fast Track kit (Invitrogen). Two micrograms of poly(A)⁺ RNA from each tissue was separated in a formaldehyde/0.8% agarose gel and transferred to a Hybond-N nylon membrane (Amersham) by capillary action. Hybridization was performed in buffer containing 50% formamide, 5 \times SSC, 20 mM Na₂HPO₄/NaH₂PO₄ (pH 7.6), 7% SDS, 1% PEG (*M_w* 20,000), and 0.5% nonfat powdered milk. The *v-qin* probe extending from nt 768 to nt 1397 (1) was used. Final washes were in 0.2 \times SSC/0.1% SDS at 55°C.

Southern Blot Analysis of CEF Genomic DNA. Ten micrograms of CEF genomic DNA was digested with *Bam*HI or with *Eco*RI. The digested DNA was then separated in a 0.8% agarose gel and transferred to a Hybond-N nylon membrane by capillary action. Hybridization conditions, probe, and washes were as described above for Northern blot analysis.

***In Situ* Hybridization.** *In situ* hybridization with whole 4-day chicken embryos was performed as described (20). For whole-mount embryos, digoxigenin-labeled RNA probes were synthesized by T7 (antisense) or T3 (sense) RNA polymerase from linearized plasmid templates containing *v-qin* DNA from nt 768 to nt 1397. Hybridization was at 70°C for 16 hr, and high-stringency washes were in 50% formamide/2 \times SSC at pH 5.8. The specificity of the hybridization signal was confirmed

Abbreviations: BF-1, brain factor 1; CEF, chicken embryo fibroblast; HNF, hepatocyte nuclear factor.

*H.W.C. and J.L. contributed equally to this work.

†Present address: Division of Biology, California Institute of Technology, Pasadena, CA 91125.

‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. L36814).

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

served region of the winged helix family. From a screen of 4×10^5 clones, 3 positive clones were obtained. One clone, designated CEQ 3-1, was 1.6 kb in length and contained the entire coding sequence of the Qin protein (Fig. 1). Three putative ATG initiation codons were identified, starting at nt 40, nt 166, and nt 175. The ATG at nt 166 of CEQ 3-1 is the first putative translation initiation site with the Kozak consensus sequence (15). The open reading frame beginning at nt 166 encodes 451 aa. The CEQ 3-1 cDNA was transcribed and translated *in vitro* with T3 RNA polymerase and rabbit reticulocyte lysate. The translated protein has a molecular mass of 46 kDa by SDS/PAGE and is recognized by an antibody against the v-Qin protein (data not shown). These results, together with an amino acid sequence comparison between the v-Qin protein and the open reading frame of CEQ 3-1, indicate that CEQ 3-1 codes for the chicken c-Qin. The differences between the deduced amino acid sequence of chicken c-Qin protein and that of its viral counterpart are as follows (Fig. 2). At the amino terminus, the v-Qin protein is fused to a virus-coded region representing part of the retroviral Gag protein. The Gag and Qin regions are connected by eight cell-coded amino acids which lie immediately upstream of the putative Qin initiator methionine. At the carboxyl terminus, the v-Qin protein is truncated by a premature stop codon due to a single base deletion. In addition, the last 8 aa of the v-Qin protein do not share homology with the c-Qin protein. Within the v-*qin* gene, there are three single nucleotide changes, which result in two nonconservative aa substitutions. One of these substitutions (Ser \rightarrow Asn) is located in a region of the v-Qin protein that is similar to one of the conserved presumptive transactivation domains of winged helix family members (16). The other substitution (Gly \rightarrow Asp) is located in a region of v-Qin which shares close homology to the DNA-binding domain of the winged helix family (2, 3).

Northern Blot Analysis Shows Expression of c-*qin* in Chicken Embryo Brains. The expression pattern of *qin* was examined by Northern blot analysis of poly(A)⁺ RNA isolated from chicken embryo tissues. Multiple mRNA species with sizes between 4.5 kb and 3.0 kb were present in samples from chicken embryo forebrain probed with the v-*qin*-specific fragment (1). Expression of *qin* was not detected in CEFs, hindbrain, liver, intestine, muscle, or heart (Fig. 3). In addition, the expression pattern of *qin* in adult chicken tissues was also examined by Northern analysis. Expression of *qin* was restricted to the forebrain of adult chickens (data not shown).

A *qin*-Specific Probe Detects a Single-Copy Gene in the Chicken Genome. DNA was extracted from CEFs and digested with restriction endonucleases. The digested genomic DNA was electrophoresed, transferred to a nylon filter, and hybridized with the *qin*-specific probe (Fig. 4). This probe detected a single restriction fragment in CEF DNA, suggesting that it is specific for a single-copy gene in chickens.

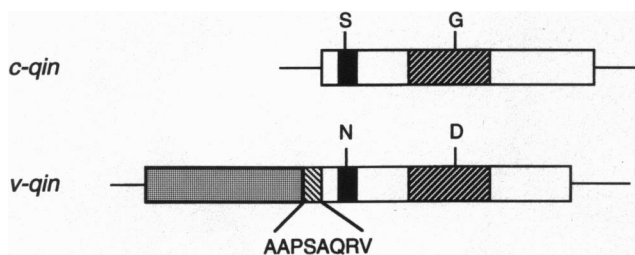


FIG. 2. Comparison of c-*qin* and v-*qin*. □, Virus-coded Gag region; ■, the eight cell-coded amino acids that are fused to the Gag region; ▨, putative transactivation domain; ▩, winged helix DNA-binding domain. Amino acids that are different between v-Qin and c-Qin are shown in the single-letter code. The lines represent untranslated sequences.

In Situ Hybridization Localizes *qin* mRNA to the Telencephalon. To determine the location of *qin* expression in developing embryos, *in situ* hybridizations to whole embryos and sections of embryos were performed. *In situ* hybridization of whole chicken embryos with a digoxigenin-labeled antisense probe for *qin* showed that the expression of *qin* was restricted to the telencephalon in 4-day embryos (Fig. 5). Examination of chicken embryo brain sections by *in situ* hybridization with an ³⁵S-labeled antisense probe for *qin* confirmed the telencephalon-restricted expression pattern (Fig. 6).

DISCUSSION

v-Qin differs from its cellular progenitor by several mutations. Whether these mutations contribute to its cell-transforming and tumor-inducing activity is not known. Preliminary results suggest that c-*qin* overexpressed by a retroviral vector can induce morphological transformation of CEFs *in vitro*.

The v-Qin protein binds to the B2 DNA site found in the promoter of the HNF-1 α gene which is the target of HNF-3; it binds less well to the HNF-3 sites in the promoter of the transthyretin gene, another target of HNF-3 (J.L., E. J. Parker, and P.K.V., unpublished results from 1993). It will now be important to determine the preferred DNA target sequences for c-Qin and v-Qin by PCR selection (17–19) in order to test

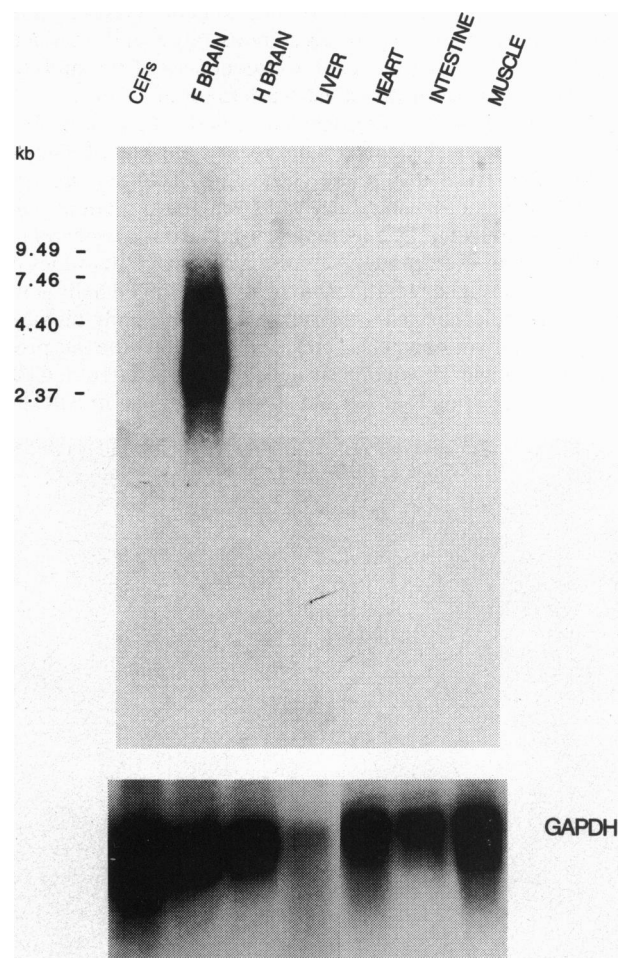


FIG. 3. Northern blot analysis of poly(A)⁺ RNA isolated from chicken embryo tissues. One major mRNA species (3.0 kb) and two minor mRNA species (4.0–4.5 kb) were detected in the chicken embryo forebrain. No signal was detected in the other tissues, even with prolonged exposure of the blot. The same blot was reprobbed with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) cDNA to standardize the quantity of RNA. F brain, forebrain; H brain, hindbrain.

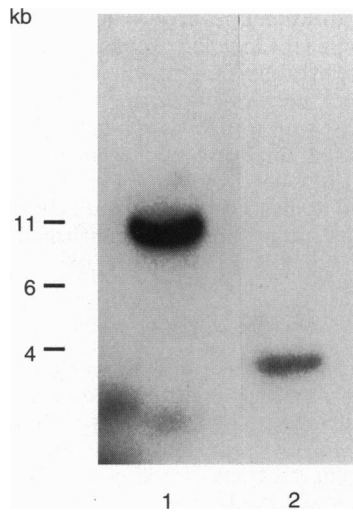


FIG. 4. Southern blot analysis of *qin* in CEF genomic DNA. *Bam*HI (lane 1) or *Eco*RI (lane 2)-digested CEF genomic DNA was electrophoresed, transferred to a nylon filter, and hybridized to the radio-labeled *qin* probe.

for possible functional consequences of the amino acid substitution in the v-Qin DNA binding domain. Transcriptional regulation by Qin has yet to be demonstrated, and here again it will be important to test for consequences of the mutation in the presumptive transactivation domain of v-Qin.

The multiple species of mRNA detected with a *qin* probe in brain tissue of 18-day chicken embryos and of adult chickens could result from differential splicing of primary *qin* transcripts. They could also be derived from genes related to *qin*; the winged helix family encompasses numerous closely related members, and even though we used a *qin* probe that does not show close sequence relationship to known winged helix genes, we cannot rule out cross hybridization with family members that have not yet been characterized. However, the *qin* probe detects a single restriction fragment in DNA from CEFs, strongly suggesting that *qin* is a single-copy gene in chickens.

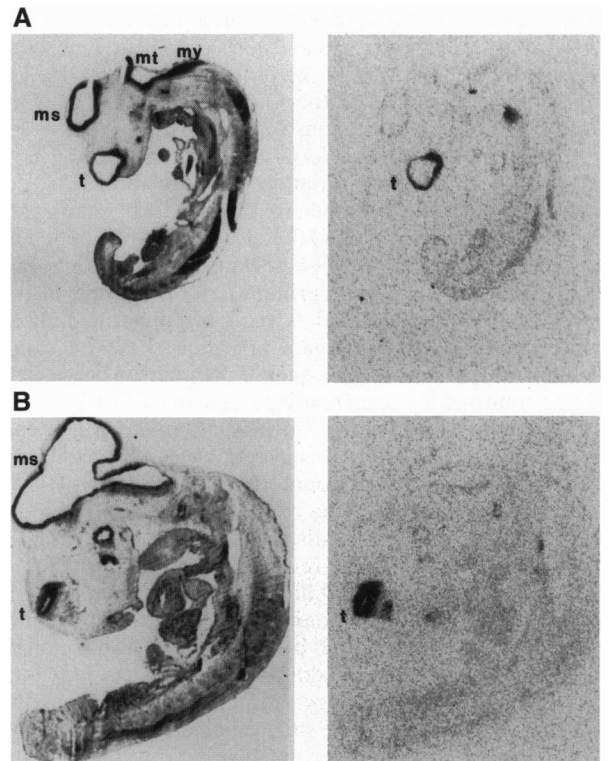


FIG. 6. Expression pattern of *c-qin* in developing chicken embryos determined by *in situ* hybridization of tissue sections. Bright-field photomicrographs (Left) and the corresponding autoradiographs (Right) of cross sections from 3-day (A) and 4-day (B) chicken embryos hybridized with a *qin* antisense probe. t, Telencephalon; ms, mesencephalon; mt, metencephalon; my, myelencephalon.

The preferential expression of *qin* in the telencephalon brings to mind the close sequence relationship between the Qin protein and rat BF-1 (1, 9). BF-1 is expressed specifically in the telencephalon of the developing rat embryo, and it is possible that Qin is the avian homolog of BF-1. However, there

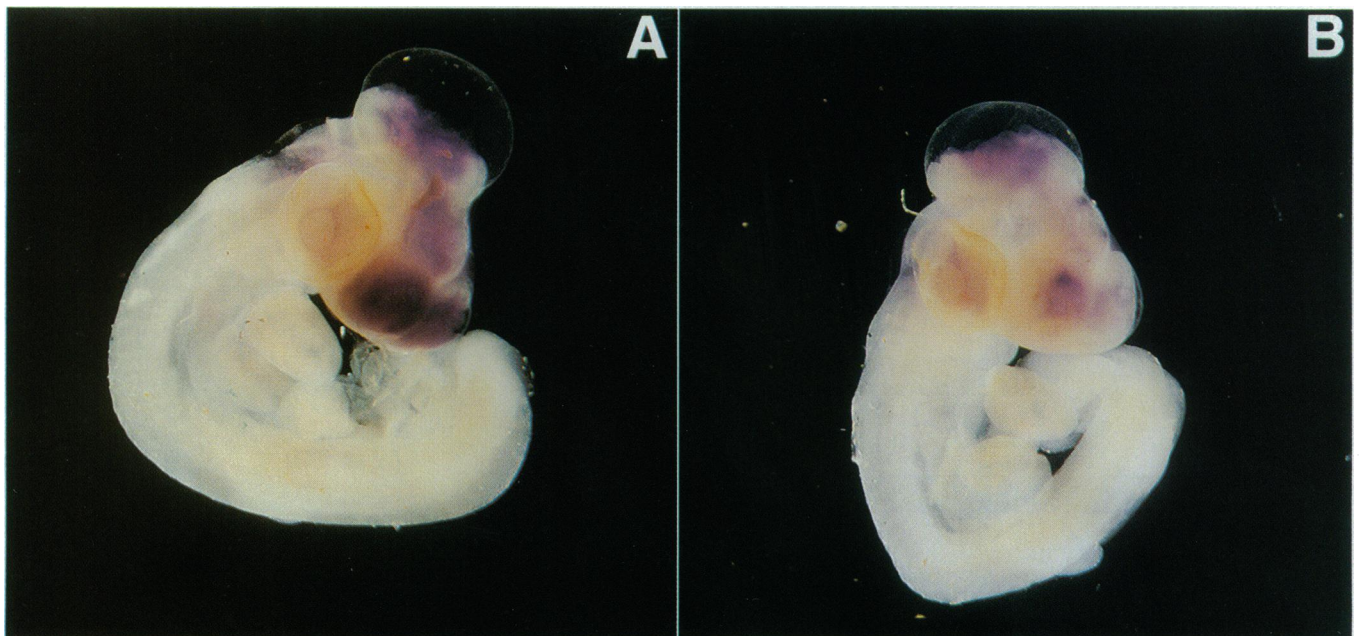


FIG. 5. Expression pattern of *c-qin* in developing chicken embryos determined by *in situ* hybridization to whole mounts. (A) Lateral view of a 4-day chicken embryo. A *qin* antisense probe was used. (B) Lateral view of a 4-day chicken embryo. A sense probe was used as control for background staining.

is significant sequence divergence between Qin and BF-1 in the amino-terminal region as well as in the untranslated regions of the mRNA that lie 5' and 3' to the reading frame. These differences may reflect the different animal phyla or different genes. The early and highly tissue-specific expression of *qin* suggests a role in brain development; the oncogenic potential of the gene may indicate a function in the control of cell growth.

We thank Cheng-Ming Chuong for providing the sections of chicken embryos. We thank Ron Evans and David Loskutoff for their helpful suggestions with the *in situ* hybridization experiments. This work was supported by United States Public Health Service Grant CA42564. H.W.C. is supported by the Cancer Research Fund of the Damon Runyon-Walter Winchell Foundation Fellowship, DRG-1304.

1. Li, J. & Vogt, P. K. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 4490–4494.
2. Weigel, D. & Jäckle, H. (1990) *Cell* **63**, 455–456.
3. Lai, E., Prezioso, V. R., Tao, W., Chen, W. S. & Darnell, J. E., Jr. (1991) *Genes Dev.* **5**, 416–427.
4. Clark, K. L., Halay, E. D., Lai, E. & Burley, S. K. (1993) *Nature (London)* **364**, 412–420.
5. Weigel, D., Jürgens, G., Küttner, F., Seifert, E. & Jäckle, H. (1989) *Cell* **57**, 645–658.
6. Grossniklaus, U., Pearson, R. K. & Gehring, W. J. (1992) *Genes Dev.* **6**, 1030–1051.
7. Lai, E., Clark, K. L., Burley, S. K. & Darnell, J. E., Jr. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 10421–10423.
8. Dirksen, M. L. & Jamrich, M. (1992) *Genes Dev.* **6**, 599–608.
9. Tao, W. & Lai, E. (1992) *Neuron* **8**, 957–966.
10. Li, C. & Tucker, P. W. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11583–11587.
11. Costa, R. H., Grayson, D. R. & Darnell, J. E., Jr. (1989) *Mol. Cell. Biol.* **9**, 1415–1425.
12. Lai, E., Prezioso, V. R., Smith, E., Litvin, O., Costa, R. H. & Darnell, J. E., Jr. (1990) *Genes Dev.* **4**, 1427–1436.
13. Keeton, M., Eguchi, Y., Sawdey, M., Ahn, C. & Loskutoff, D. J. (1993) *Am. J. Pathol.* **142**, 59–70.
14. Coligan, J. E., Kruisbeek, A. M., Margulies, D. H., Shevach, E. M. & Strober, W. (1990) *Current Protocols in Immunology* (Wiley, New York), pp. 9.4.1–9.4.3.
15. Kozak, M. (1986) *Cell* **44**, 283–292.
16. Pani, L., Overdier, D. G., Porcella, A., Qian, X., Lai, E. & Costa, R. H. (1992) *Mol. Cell. Biol.* **12**, 3723–3732.
17. Blackwell, T. K. & Weintraub, H. (1990) *Science* **250**, 1104–1110.
18. Pollock, R. & Treisman, R. (1990) *Nucleic Acids Res.* **18**, 6197–6204.
19. Thiesen, H. J. & Back, C. (1990) *Nucleic Acids Res.* **18**, 3203–3209.
20. Wilkinson, D. G. (1992) *In Situ Hybridization* (IRL, Oxford), pp. 75–83.