**RIN ZF**, a Novel Zinc Finger Gene, Encodes Proteins That Bind to the CACC Element of the Gastrin Promoter*

(Received for publication, November 10, 1998, and in revised form, January 8, 1998)

Loyal G. Tillotson‡

From the Division of Digestive Diseases and Nutrition, Department of Medicine, University of North Carolina, Chapel Hill, North Carolina 27599

Expression of gastrin, a gut hormone and growth factor, has tissue-specific transcriptional regulation and can be induced in some tumors. Previous studies have shown that a CACC cis-regulatory element is important for transcriptional activation in pancreatic insulinoma cells. To identify CACC-binding proteins, a λ phage cDNA library derived from a rat insulinoma cell line, RIN 38A, was screened by a Southwestern method. A novel member of the Cys2-His2 zinc finger gene family was cloned and designated RIN ZF, having a cDNA sequence of 3.8 kilobases. One full-length and a shorter splice variant were sequenced and had predicted protein masses of 91.6 and 88.7 kDa. Expression of both splice forms were ubiquitous in fetal and adult rat tissues. Recombinant RIN ZF protein exhibited sequence-specific binding to the gastrin CACC element in a gel mobility shift assay. In transient transfections, both splice variants appeared to have only weak activating effects on gastrin-luciferase reporter gene transcription. Furthermore, RIN ZF coexpression with Sp1 appeared to block the strongly activating effects of Sp1 mediated through the CACC element. These findings suggest that a novel set of zinc finger proteins may help regulate gastrin gene expression by interfering with Sp1 transactivation.

Gastrin, a peptide hormone and growth factor of the gastrointestinal tract, is expressed in a developmental and tissue-specific manner. In development, gastrin is highly expressed in fetal pancreatic islet precursor cells and in the fetal colon (1, 2). After birth, expression greatly diminishes in those sites and shifts to expression in the stomach and proximal small intestine. Gastrin expression occurs in adenoacarcinoma of the colon (3) and pancreas (4) and in some pancreatic endocrine tumors. In these neoplastic tissues, gastrin functions as a tumor growth factor. Therefore, elucidation of transcriptional control mechanisms may provide insights into induction of oncofetal gene expression in neoplasia.

Transgenic studies with rat-human gastrin gene hybrids demonstrate appropriate tissue-specific and developmental expression as long as the promoter contains proximal 5'-flanking sequence joined to the first exon (5). Further studies have shown that gastrin transcriptional regulation is in large part mediated through trans-factor interactions along the proximal 5'-flanking sequence (5, 6). Multiple cis-regulatory elements have been identified in the gastrin promoter, including Sp1 sites (7, 8), homeodomain-like elements (8), a negative element (9), E-box binding sites (10), an epidermal growth factor-regulated protein binding site (11), and a CACC element (6).

Systematic scanning mutagenesis of the proximal gastrin promoter suggests that the CACC element, located approximately 110 bp from the transcriptional start site, is necessary for promoter activity in rat insulinoma (8) and human colon cancer cell lines (12). Mutational analysis of the CACC element showed a close correlation between DNA binding by protein complexes and transcriptional activation of gastrin-luciferase reporter genes in insulinoma cell lines. Efforts to purify DNA-binding proteins by sequence-specific affinity chromatography (13) were successful, but the yield was too low for peptide sequencing (6). Therefore, a different approach was followed by probing electrophoretic blots of insulinoma cell extracts with multimerized CACC element DNA probes. These Southwestern blots revealed several DNA-binding protein bands from 70–110 kDa in size (6). Sequence specificity of the DNA binding was verified by probing with a multimerized analog of the CACC element containing two base mutations. The same mutations rendered a gastrin-luciferase reporter gene functionally inactive, and hence, when used as a southwestern probe, clearly differentiated specific from nonspecific binding. In the current study, a Southwestern binding approach was employed to screen a λ phage, insulinoma cDNA expression library for CACC-binding proteins.

**EXPERIMENTAL PROCEDURES**

cDNA Library Construction and Screening—RIN 38A cell total RNA preparation and two rounds of poly(A) selection (oligo(dT)-cellulose spin column; Amersham Pharmacia Biotech) were accomplished by established methods (14). A λ gt11 cDNA library was then commercially prepared (Stratagene, La Jolla, CA). Subsequent screening of the library for CACC DNA binding employed the Southwestern blot method of Singh and co-workers (15). The unamplified library (10^6 plaque-forming units) was plated with Y1090 bacteria and induced with isopropyl-1-thio-galactoside-coated nitrocellulose filters for 6 h. Filters were then treated and probed as described previously (6). Nick-translated, double-stranded DNA probes were prepared from concatamers of CACC element containing sequences (CCCCCCCCCAT)₆ (wild type minimal CACC element), (CCCCCCCCATCCCCTCTGCCTGGA)₄ (wild type), and (CCCCACCATTTCTCTGGCCTTGGAA)₄ (mutant gastrin promoter elements). After four rounds of plaque isolation, those phage positive for wild type binding and negative for the mutant probe were amplified and prepared (lambda kit, Qiagen, Valencia, CA). Phage cDNA inserts were subcloned into pGem 7 vectors (Promega, Madison, WI) and sequenced by dyeoxy sequencing methods (Sequenase; Amer-

---

*This work was supported by National Institutes of Health Grant R29 DK49860. 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(s) reported in this paper has been submitted to the GenBank®/EJB Data Bank with accession number(s) AF091457.

† To whom correspondence should be addressed: CB7038, Rm. 148 Glaxo Bldg., University of North Carolina, Chapel Hill, NC 27599-7038. Tel.: 919-966-7884; Fax: 919-966-7468; E-mail: lgtgl@med.unc.edu.

1 The abbreviations used are: bp, base pair(s); RT-PCR, reverse transcriptase polymerase chain reaction; GST, glutathione S-transferase; ORF, open reading frame; RIN ZF, zinc finger gene cloned from rat insulinoma 38A cells; RIN ZFsv, splice variant; BLAST, basic local alignment search tool.
sham Pharmacia Biotech). Subsequent rounds of library screening for cDNA fragments overlapping with clone 25A were accomplished by nucleic acid hybridization with nick-translated fragments of EcoRI/SalI 25A and EcoRI/BamHI 25B digests. Clone 25F had approximately 75% GC content, which necessitated sequencing by the Maxam-Gilbert chemical method (14). All cDNA fragments subcloned in pGEM 7 were sequenced along both sense and antisense strands to confirm the sequence information. Sequence formatting, assembly, and analysis were aided by MacVector and AssemblyLIGN programs (Oxford Molecular Group, PLC) and GenBank 228 (National Center for Biotechnology Information).

**RIN ZF Cloning—**A 5'-truncated version of RIN ZFsv was developed by PCR amplification from clone 25C of a 962-bp fragment (bases 1181–2143) that was subsequently ligated into the SalI site of clone 25A-pGEM 7. The recombinant product was then excised with BamHI digest and ligated into a pGex 2T bacterial expression vector (Amersham Pharmacia Biotech). Recombinant glutathione S-transferase (GST), GST-truncated RIN ZFsv, and GST-Elk were expressed in JM109 Escherichia coli by a modified bacterial expression method (17). Briefly, transformed bacteria were grown at 37 °C overnight (50 μg/ml ampicillin). The bacterial colonies were then suspended into LB broth (100 μg/ml ampicillin, 0.2 mM isopropyl-1-thio-β-D-galactoside, 10 μM ZnSO4) and incubated at 30 °C for 3 h (200 rpm shaking) before harvest and protein extraction. Extracts were prepared by sonication and purified by glutathione-Sepharose affinity binding by well established methods (18). Gel mobility shift assays were by previously described methods (6).

**RESULTS**

RIN ZF Binds Gastrin CACC Element

The full-length coding region and 3'-untranslated sequences of RIN ZF and RIN ZFsv were excised and ligated in-frame with a start codon into an actin promoter-driven expression vector, pPac (19). The coding region and 3'-untranslated region for clone 25A, designated truncated RIN ZF (Fig. 5A), was ligated in-frame with a start site into a pPac expression vector. Expression of all three versions was confirmed by a luciferase reporter plasmid (6) included one with a proximal gastrin promoter and another with a concatamer of the human gastrin CACC element. Twelve positive clones were identified and subsequently screened with a mutated CACC concatamer probe to verify sequence specificity. Four true-positive CACC-binding phage isolates were then plaque-purified. After subcloning and sequencing the cDNA inserts, gene fragments homologous to rat Sp1 (accession number D12768) (20) and closely homologous to mouse Pur-1 (95% identity, accession number L04649) (21) and human SPR-2 (94% identity, accession number X68560) (22) were identified. All of the cDNA inserts identified in this assay for CACC binding possessed zinc finger domains of the Cys2-His2 type, including one, clone 25A, which otherwise had no close homology with known genes. The latter clone was chosen for further investigation.

Clone 25A was fully sequenced and found to be 2082 bases long and had an open reading frame (ORF) from the 5' end extending 1413 bases. Within the 25A ORF were two complete and one partial Cys2-His2 zinc finger domains. The subsequent screening, plaque purification, subcloning, and sequencing of cDNA fragments homologous to 25A is shown in schematic form in Fig. 1. cDNA fragments 25B, C, D, and E were identified by screening for homology to the 5'-Sal fragment of 25A. Clones 25B and E were identical in their overlapping regions, as were 25C and D. However, the pairs differed from each other presumably by splice variation in an 87-base pair region that was not present in 25C and D. A 5'-BamHI fragment from 25B was then used to screen the library for the clone 25F. The overlapping sequences of 25F with 25B–E were identical. As shown in Fig. 2, assembly of the overlapping 25A–F-cloned fragments yielded a cDNA sequence of 3850 bases.

Analysis of the assembled cDNA sequence revealed a single long ORF from 674 to 3180 (Fig. 2). The ATG codon from base 674 is within a favorable context for initiation (23). This putative initiator codon has three upstream, in-frame terminator codons. Of the other three upstream ATG sequences present, all are out-of-frame and have a very weak context for initiation. The full-length ORF encodes a putative protein of 836 amino acids in length with a predicted molecular weight of 91 kDa. As previously noted, several of the cDNA sequences differed in the lack of an 87-base pair sequence shown in Fig. 2 in the boxed sequence. This apparent splice variation is located between bases 1457 and 1544, which is equivalent to a 29-codon difference. The exon donor and recipient half-sites conform with known splice site consensus motifs (24). The predicted molecular weight of the smaller splice variant is 89 kDa. The zinc finger motifs are of the Cys2-His2 type and are located near the COOH-terminal region. Two complete zinc finger domains and one partial domain are present and have the typical seven-codon HC connector sequences. However, these sequences are not similar to the Krüppel consensus motifs (25). Gapped BLAST (V2.0) (16) sequence analysis of the RIN ZF DNA and
its translated sequences did not find any matches or close homologies with known genes.

**RIN ZF mRNA Expression**—To determine the expression of the RIN ZF genes, poly(A)-selected RNA from Rin 38A, Rin B6, and Rin 1056 cell lines were probed in Northern blot analysis. Only three weak bands were detected at approximately 2.0, 5.0, and 9.0 kb in size (data not shown). An RT-PCR approach was then employed to determine the tissue and developmental expression pattern of the RIN ZF genes. Oligonucleotide primers flanking the putative splice sites were used in the RT-PCR assay shown in Fig. 3. Panels of adult and fetal rat RNA samples were screened and yielded the predicted PCR products of 791 and 704 bp in size. Both PCR products were isolated, subcloned, and sequenced, confirming their identity with the selected regions of the full-length and splice variant cDNA library clones. The RIN ZF gene appears to be ubiquitously expressed, having message present in all the fetal and adult rat tissues tested. Although both forms of RIN ZF were detected, the relative expression of each may be variable because this RT-PCR assay was not quantitative. RNA from RIN cell lines 38A, B2, B6, 1056C, rat fibroblast Rat2, myoblast L8, and pancreatic AR42J cell lines also yielded PCR products of 791 and 704 bp (data not shown).

**Recombinant RIN ZF Gel Mobility Shift Assay**—An expression vector for GST fusion-RIN ZF protein was developed by ligating an in-frame sequence of a truncated form of the splice variant into pGEX 2T vector. Recombinant GST-RIN ZF was prepared and had an approximate molecular size of 66 kDa (data not shown). Recombinant GST, GST-Elk, and GST-RIN ZF were used in a gel mobility shift assay with a gastrin CACC element probe. As shown in Fig. 4, the GST-RIN ZF binding was sequence-specific, because a 50-fold excess of cold wild type competitor oligonucleotide completely competed probe binding. However, a 50-fold excess of an oligonucleotide with a 2-bp mutation failed to compete. Therefore, recombinant RIN ZF

---

**FIG. 2.** RIN ZF cDNA sequence with open reading frame translation. Boxed bases identify the splice variation. Double-dotted underlined bases denote the zinc finger domains. Single and double-underlined bases denote forward and reverse PCR primer sites, respectively. The translated sequence of RIN ZF ORF is noted in italics.

**FIG. 3.** Reverse transcriptase PCR amplification of RIN ZF primed mRNA from rat tissues. Total RNA was isolated from fetal and adult rat tissues and used in RT-PCR reaction as noted under "Experimental Procedures." RIN 38A cell total RNA is denoted as 38A, with (+RT) or without (−RT) reverse transcriptase enzyme. RT-PCR products and DNA size markers were resolved on 1.5% agarose gel. RIN ZF primers encompassed the splice variation region as noted in Fig. 2.
has the correct sequence specificity of DNA-binding proteins that interact with the gastrin CACC cis-regulatory element. It is noteworthy that the doublet banding pattern of the gel-shifted probe obtained with GST-RIN ZF is similar to that previously observed with Rin 38A nuclear extracts (6).

**RIN ZF Activation of Gastrin-Luciferase Expression**—It was initially predicted that RIN ZF would be an activating transcription factor because the CACC element was a positive cis-regulatory element in the gastrin promoter. However, initial results of expression of RIN ZF in mammalian cells showed only weak activation, perhaps because of the presence of other transcription factors. Therefore, RIN ZF expression vectors were transfected into Drosophila SL2 cells, as they had no Sp1-like activity (26). As shown in Fig. 5A, pPac vectors for the expression of full-length, splice variant, and 5'-truncated forms of RIN ZF were utilized. These expression vectors were cotransfected into SL2 cells with a proximal gastrin promoter luciferase reporter gene, and the results shown in Fig. 5B. The full-length and splice variant versions had only modest activating effects on 200 Gastrin-Luc reporter gene expression, resulting in 2–3-fold increase over the empty vector, pPac. A 5'-truncated version of RIN ZF lacking 365 codons of the 5'-coding region but still possessing the zinc finger DNA binding domains failed to activate the reporter gene expression. Because the function of some zinc finger transcription factors may be concentration-dependent (28), RIN ZF expression plasmids were transfected over a concentration range of 0.05–2.0 μg/well. However, there was no significant change in the overall pattern of weak transcriptional activation of gastrin-luciferase expression (data not shown).

To examine whether RIN ZFsv can act synergistically with Sp1 or inhibit Sp1 transactivation, expression plasmids for both were cotransfected into SL2 cells with a (CACC)6pT81Luc reporter gene. As shown in Fig. 6, Sp1 strongly transactivated CACC-Luc expression compared with the modest transactivation associated with RIN ZFsv. When the two expression vectors were cotransfected in equivalent amounts, the CACC-Luc transactivation was 80% lower than the Sp1 alone. Increasing the amounts of RIN ZFsv further diminished but did not completely block the Sp1 effect. Similar results were obtained from cotransfection of the full-length form of RIN ZF with Sp1 (data not shown). These results suggest that RIN ZF does not act synergistically with Sp1 but rather competes for binding to the CACC element.

**DISCUSSION**

Several hundred transcription factors have been cloned and identified by a variety of methods over the past decade (27). Some, like Sp1, were identified after purification of the native factor by DNA affinity chromatography (13). Although this approach was initially applied to the gastrin CACC-binding
protein (6), it proved to be impractical given the low abundance of the CACC binding factors. An alternative approach of screening of a cDNA expression library by DNA binding (28) was pursued because several CACC-binding proteins ranging in size from 70–110 kDa were evident on a Southwestern blot of RIN 38A cell extracts (6). Potential disadvantages of this approach are that many transcription factors do not retain their DNA binding properties under the screening conditions, and some require co-factors or dimerization to bind in a sequence-specific manner. In addition, DNA-binding proteins possessing broad sequence specificity can generate false positive clones (29). With these caveats in mind, screening a RIN 38A cDNA library yielded several clones that showed strong binding activity to a concatamer of the gastrin CACC element. A secondary screen with mutant CACC concatamer probe verified sequence specificity by RIN ZF and several known zinc finger transcription factors. To ensure that RIN ZF CACC binding was not an artifact of a concatamered probe, sequence specificity was confirmed by gel mobility shift assay with a monomeric CACC element probe.

As previously noted, although several CACC binding proteins were detected in RIN cell nuclear extracts, a 70-kDa protein appeared most prominent. Interestingly, none of the positive clones identified in the Southwestern screen appeared to encode a 70-kDa protein. Although unlikely, it is possible that RIN ZF and the 70-kDa protein are the same if RIN ZF has an alternative translational initiation site from the one predicted by ORF analysis. A number of viral and cellular genes have been identified that yield multiple protein products resulting from alternative initiation of translation. It has been observed that growth factor and transcription factor mRNAs possessing long G + C-rich 5′-untranslated regions exhibit alternative translation initiation, often through internal ribosome entry sites (30–32). Similarly, RIN ZF has greater than 75% G + C content in the 5′ kilobases of sequence, which likely has significant secondary structure inhibitory of ribosomal scanning. The issue of the relative size of native RIN ZF proteins remains to be resolved by immunoblotting studies with antisera prepared against recombinant RIN ZF.

Apart from the Cys2-His2 zinc finger domains, the RIN ZF coding sequence bears little resemblance to other reported zinc finger genes. Specifically, there were no Krüppel (25) or BTB domains (33), which were first described in Drosophila but are common in higher organisms. Other transcription factor domains such as homeotic (34, 35), leucine zipper (36, 37), Ets (38), and basic helix-loop-helix (39) also were not found in RIN ZF. Subdividing the RIN ZF coding sequence and searching by BLAST alignment failed to reveal any close homology with known genes. However, searching an expressed sequence-tagged data base revealed a close match with an expressed sequence tag derived from microdissected human prostate intraepithelial neoplasia 2 cells (accession number AA602975). This 299-bp expressed sequence tag overlapped 221 bases of the RIN ZF from 2041 to 2265, with an homology of 94%. In separate studies, we used the RIN ZF PCR primers and DNA prepared from a human gastric tumor cell line, AGS, to amplify the RIN ZF from 2044 to 2265, with an homology of 94%. In this 299-bp expressed sequence tag overlapped 221 bases of the RIN ZF from 2041 to 2265, with an homology of 94%. In separate studies, we used the RIN ZF PCR primers and DNA prepared from a human gastric tumor cell line, AGS, to amplify the RIN ZF from 2044 to 2265, with an homology of 94%. In this 299-bp expressed sequence tag overlapped 221 bases of the RIN ZF from 2041 to 2265, with an homology of 94%. In separate studies, we used the RIN ZF PCR primers and DNA prepared from a human gastric tumor cell line, AGS, to amplify the RIN ZF from 2044 to 2265, with an homology of 94%. In this 299-bp expressed sequence tag overlapped 221 bases of the RIN ZF from 2041 to 2265, with an homology of 94%.

The mRNA expression pattern of RIN ZF is not limited to those tissues that express gastrin. Both splice forms appear to be ubiquitously expressed and not developmentally regulated. However, these observations are limited by the RT-PCR method to a qualitative assessment of message in whole tissues. Immunohistochemistry with antisera specific to the full-length and splice variant forms of RIN ZF would better define the expression pattern and perhaps provide insights into functionality. It is noteworthy that some zinc finger transcription factors are developmentally regulated by expression of alternatively spliced variants (40). Several transcription factors were found to have splice variations in the DNA binding domains (41) as well as in activation or repression domains (42). However, the RIN ZF splice site domain, which is 5′ to the zinc finger domains, consists of largely neutral amino acids and has no recognizable motif. Therefore the functional differences between the splice variants remains to be determined.

Both full-length and splice variant RIN ZF appeared to have weak trans-activating effects on the gastrin promoter, whereas the truncated form had no effect (Fig. 5B). These results indicate that within the RIN ZF amino-terminal region is an activation domain, despite the lack of recognizable activation motifs, such as homopolymeric glutamine and proline-rich stretches (43). Instead, RIN ZF possesses several regions rich in basic residues, often associated with repressor domains (44). It may be that RIN ZF have intrinsically weak transactivation domains or may require the presence of coactivating factors (45) not present in the Drosophila cells. Some zinc finger proteins contain both activation and repression domains (46) and may be regulated by the context of the promoter binding site or transcriptional cofactors. Further studies with chimeric RIN ZF deletional constructs are in progress to determine functional activity of RIN ZF domains.

It has been shown that some zinc finger proteins, such as Krüppel, have opposite regulatory effects that are concentration-dependent (47). Sauer and Jäckle (47) found that monomers of Krüppel activate, whereas dimers repress transcription even though both forms bind to the same element. In contrast, RIN ZF expression vector did not vary in its activating effects on gastrin-luciferase transcription throughout a wide concentration range. Alternatively, RIN ZF may have a down-regulatory effect on gastrin transcription by interfering with Sp1 binding or function. BKLF, or basic Krüppel-like factor, is one example of a zinc finger protein that competes with Sp1 binding at a CACCC element of the β-globin gene in erythroid cells (48). Similarly, ZBP-89 competes with Sp1 in binding to a GC-rich proximal element in the gastrin gene (49). As Merchant and her colleagues have shown (49), ZBP-89 inhibits Sp1 binding and blocks epidermal growth factor induction of gastrin transcription. Other mechanisms of altering Sp1 activity include functional interactions with other transcription factors (50), changing the Sp1 phosphorylation state (51) or targeting Sp1 for proteolysis (52). The most likely mechanism is that RIN ZF competes with Sp1 binding at the CACC element, but further studies are needed to determine binding affinities and footprints.

Acknowledgments—I am grateful to Philip J. Davis, James F. Preston, and Dr. Shu-Li Wang for technical assistance. Many thanks to Dr. Todd S. Gulick for advice regarding recombinant protein expression. I greatly appreciate the support and encouragement given by Dr. Stephen J. Brand at the start of this project.

REFERENCES
1. Larsson, L., Rehfeld, J., Sundler, F., and Hakanson, R. (1976) Nature 262, 609–610.
2. Luttichau, H. R., Van Solinge, W. W., Nielsen, F. C., and Rehfeld, J. F. (1993) Gastroenterology 104, 1092–1098.
3. Van Solinge, W. W., Nielsen, F. C., Friis-Hansen, L., Fulkmer, U. G., and Rehfeld, J. F. (1993) Gastroenterology 104, 1099–1107.
4. Smith, J. P., Fantasky, A. P., Liu, G., and Zagon, I. S. (1995) Am. J. Physiol. 268, R135–R141.
5. Wang, T. C., Babayatsky, M. W., Oates, P. S., Zhang, Z., Tillotson, L., Chulak, M., Brand, S. J., and Schmidt, E. V. (1995) Am. J. Physiol. 268, G1925–G1936.
6. Tillotson, L. G., Wang, T. C., and Brand, S. J. (1994) J. Biol. Chem. 269, 2234–2240.
7. Merchant, J. L., Demediuk, B., and Brand, S. J. (1991) Mol. Cell. Biol. 11,
